Texts and Monographs in
Economics and Mathematical Systems

Edited by
Martin J. Beckmann and Wilhelm Krelle

Martin J. Beckmann

Tinbergen Lectures on Organization Theory

With a Preface by Jan Tinbergen

With 15 Figures

Springer-Verlag
Berlin Heidelberg New York Tokyo 1983

Prof. Dr. Martin J. Beckmann
Brown University, Providence, RI 02912, USA
and
Institut für Angewandte Mathematik, Technische Universität München
Barerstraße 23, D-8000 München 2, FRG

ISBN 3-540-12646-5 Springer-Verlag Berlin Heidelberg New York Tokyo
ISBN 0-387-12646-5 Springer-Verlag New York Heidelberg Berlin Tokyo

Library of Congress Cataloging in Publication Data. Beckmann, Martin J. · Tinbergen lectures on organization theory. (Texts and monographs in economics and mathematical systems). Includes bibliographical references and index. 1. Organization-Addresses, essays, lectures. I. Title. II. Series. HD31.B3699542 1983 658.4'013 83-14765
ISBN 0-387-12646-5 (U.S.)

Printing: Weihert-Druck GmbH, Darmstadt. Binding: J. Schäffer OHG, Grünstadt
2142/3140-543210

Preface

In this book Professor Beckmann, with considerable ingenuity, offers a mathematical analysis of productive organizations in the widest sense. Starting with descriptive features he builds up, step by step, production functions, profiting from the rigor of a set of axioms or assumptions and their logical implications. Among the organizations studied hierarchies play a predominant role and are compared with such forms of cooperation as partnerships and "ladders". A number of well-known basic concepts such as span of control, rank, line vs. staff and others serve as starting points. His analysis leads to such refinements as balanced, regular or degenerated organization patterns and interesting comparisons of the efficiency of various structures.

Empirical verification of the axioms or assumptions is not the objective chosen by the author--except a few concrete illustrations--but the book constitutes an excellent basis for such research.

Several of the results obtained take simpler forms for very large hierarchies. The renewed interest, shown in political discussions, in the bureaucratization of both large enterprises and government machinery makes Dr. Beckmann's work highly topical. Discussions (by Bahro) of the GDR and by many other authors of Japanese management as compared with American or Western European are cases in point. Some additional variables may then have to be added, of a psychological nature: for instance satisfaction from work or irritation evoked by excessive supervision.

Having participated in some recent empirical work about the production functions for large entities and about managers' incomes I found Professor Beckmann's coherent and imaginative theoretical setup very helpful. I am certain my evaluation will be shared by those active in related areas of economic research.

Jan Tinbergen
Emeritus Professor
of Development Programming
Erasmus University, Rotterdam

Acknowledgements

While holding the Tinbergen chair at Erasmus Universiteit Rotterdam in April, May and June of 1982, I gave two courses of lectures: one on selected problems in spatial economics, and the present one on the economics of organizations.

Encouraged by my good friend Jean Paelinck I organized the material of this as a monograph. This task was made simple by the fact that I had the material for the various chapters available in a preliminary form as papers written under NSF Grant 79-19376. In a few places I have drawn on my "Rank in Organizations" to fill in gaps.

In fact, this is a revision of the material in Chapters I, II, III, VIII, IX, and X of my earlier book, Rank in Organizations. It addresses the structural aspects of organizations as distinct from those concerned with the movement of personnel. The latter will be reconsidered in a separate monograph.

It is my pleasure to thank Erasmus Universiteit for the honor of appointing me to the Tinbergen chair in April, May and June of 1982, and the Technical University of Munich for permission to accept the appointment. I also wish to thank my faithful audience for their encouragement, interest and stimulating discussion, in particular Professor Jean Paelinck, Dr. Jean-Pierre Ancot, Dr. Thijs ten Raa, Dr. Hans Kuiper, Dr. Henk Haverstein, and Dr. Paul Beije.

In making this substantial revision and--I hope--improvement I have benefitted enormously from teaching this subject in Economics 165 "Rank in Organizations" at Brown University. The research required was generously supported by NSF Grant SES 79-19376, for which grateful acknowledgement is made. In particular, the following efforts were undertaken under this grant.

1. Develop a description of organizational structure in terms of directed graphs. Explore their structural properties.

2. Recognize the communication aspects of organizations and describe them in terms of nondirected graphs.

3. Characterize the alternative complete orderings that may be obtained from a given pre-ordering of positions (by supervision) in an organization.

4. Define and construct efficient organizational designs under a given constant span of control.

5. Explore the Uses of Production Functions in Organizational Analysis.

I should like to thank the National Science Foundation for its generous support of this effort through Grant SES 79-19376. It is a pleasure to thank Mrs. Marion Wathey for her beautiful typing of a very messy manuscript, Dr.J. Fischer of the Technical University Munich for making the drawings, and Dr.H. Mittermeier for preparing the index.

Providence, Rhode Island

April, 1983

Table of Contents

	Preface by Jan Tinbergen	V
	Acknowledgements	VII
	Notation	XI
I.	RANK	1
	1. Introduction	1
	2. Supervision	4
	3. Control	8
	4. Rank	11
	Supplement A: A Graph Theoretic Model of Supervision	19
	Supplement B: Precedence	30
	Supplement C: Communication Aspects of Rank	32
II.	DESIGN	35
	5. Span of Control	35
	6. Estimation	41
	7. Assignment	46
Supplement	D: Integers and Slack	49
	E: Average Span and Management Graph Theory	53
	F: Support Structure	58
	G: Maximizing Average Rank	60
III.	COSTS AND SCALE	64
	8. Regular Organization	64
	9. Costs and Scale	77
	10. Loss of Communication and Control	89
IV.	ORGANIZATIONAL PRODUCTION FUNCTIONS	94
	11. Case Working Organizations	94

12. A Production Function for Management 109
13. Allocation of Inputs: Cost Functions 115

V. ADVANTAGE OF ORGANIZATION 127

14. Organization versus the Individual 127
15. Supervision versus Partnership 135
16. The Economics of Hierarchy 151

Supplement H: Qualification and Specialization 166
References 168
Index 171

Notation

(Symbols used in more than one chapter)

α	output elasticity of labor
a	productivity coefficient salary factor
β	output elasticity of supervision
$\bar{c}$	unit cost
$F(.,.)$	production function
$f(\cdot,\cdot)$	linear homogeneous production function
$\phi(x)$	single factor production function
ρ	factor of control loss
g	surplus of a partnership
h^r	surplus under management by delegation r
k	surplus under team management
n_r	number of positions of rank r
N_r	number of positions of rank r or higher
N	total number of positions in the organization
π	product price
q	measure of ability
Q	size of the task, number of operatives required
r	rank
R	presidential rank, also scale of the organization
w_r	wage in rank r
w_R^*	presidential wage
W	wage bill
$\bar{w}$	average wage
x_r	positions in rank r
y_r	output of managers, "management" in efficiency units
z_r	staff positions in rank r

I. Rank

1. Introduction

The economics of organization is a subject of some current interest. There are various approaches of which this is only one. The subject has no well-defined boundaries, and this book explores only a small part of it. It is more restricted than my earlier "Rank in Organizations" by excluding the economics of careers in organizations. Its main topics are the internal structure of organizations and such questions as returns to scale, loss of control, and the economic advantage of organizations. The table of contents gives some indication of the various topics pursued. While not exhaustive this is, I believe, a coherent and self-contained treatment of some basic questions that economic theory might ask of organizations.

For purposes of this monograph organizations are defined as an association of persons using common resources to achieve well-defined but limited ends. In the literature such organizations are also referred to as "work organizations". As will become apparent, multi-level organizations have all the characteristics of Weberian bureaucracy [Max Weber, 1921].

A family, clan or nation is thus not an organization (their ends being too broad) nor is an industry (resources are not held in common) nor is a one-person firm. But the concept is broad enough to include business firms as well as nonprofit organizations with economic goals, government or self-contained parts thereof, churches, universities and last, not least, military organizations.

In order to introduce the other basic concept, rank, we must first consider supervision and the need for it (Section 2).

Although organizations are usually thought of as the preserve of sociologists, they are a potential subject for economic analysis because

i. they produce services and use resources;
ii. they may achieve a given task more effectively and efficiently than individuals on their own;
iii. they are subject to economies and/or diseconomies of scale, and this is relevant to an understanding of the optimal size of firms;
iv. they affect the structure of labor markets.

Last not least they are an important part of modern life and for that reason alone a worthy subject of economic inquiry.

The focus in this book is on supervision and its structural implications. The structure of supervision is examined in Chapter I first in terms of directed graphs, then as the basis of a pre-ordering and a simple ordering generating ranks. This scheme is to be compared to one that would result from communication--the distance relationship between members in an organization. Ideally the two structures are not in conflict but are coincident. Superior and inferior--or efficient and inefficient--structures may be distinguished already at this point.

However, more probing analysis requires that a limit on effective supervision be recognized. This is done in Chapter II by means of the well-established concept of the span of control, seen here as an upper bound on the number of persons that can be effectively supervised by one supervisor. The standard case is that this bound is the same throughout the organization. Relaxations of this far-reaching assumption are considered subsequently. Efficient and inefficient designs may now be compared, and an algorithm developed constructing an efficient design in the case that departmental boundaries are sufficiently flexible. Chapter III reconsiders the important issue of returns to scale in organizations. We consider first the cost advantage of an ideal type of organization, a regular organization, with constant span of control and delegation of all operative work to

rank zero. A regular organization has asymptotically constant unit costs and has asymptotically constant returns to scale. This is also true of similar types of organization (extended regular and quasi-regular) for which the number of operatives rises as fast as the total number of organization members. But for other types of organization this is not true. While increasing returns in administration are inconsistent with constant spans of control, decreasing returns can be generated in many ways, for instance, through a proliferation of staff positions.

But spans of control cannot be assumed as given, they are the results of economic choice. This point is made in Chapter IV first by working out the example of an organization charged with case work: supervisors and operatives are substitutable and this fact may be described in terms of the familiar concept of a production function. These production functions can then be used to re-examine the relationship between output and costs and to construct cost functions.

Chapter V finally applies production functions to a study of the important question: when and why organizations are superior to individual effort, and when a hierarchical form is preferred to an unsupervised partnership. When simple organizations (organizations with only one supervisor) are advantageous, what can be said about the advantage of hierarchies composed of many levels? The economics of such hierarchies is the subject of the final section.

We also consider briefly how personnel of different qualifications should be allocated, and why the superior ones should be appointed managers.

2. Supervision

Organizations are charged with tasks that exceed the capacity of a single individual. To get the task done, the individual contributions must be coordinated. In the absence of markets this is done through an assignment of accountability for results. But in order to get the results, to get the job done, persons who are accountable must also have authority. Thus whoever performs part of the job is made to report to the person who is in charge and held accountable for the performance of this job. We call this relationship supervision and interpret it broadly.

Thus the supervisor too must report to a supervisor, and so on. The only exception to this rule is to the top person in the organization who although accountable for the performance of the entire organization, reports to no one: the president. An organization with supervision is called a hierarchical organization.

These notions are formalized in the following abstract description of supervision in a hierarchical organization. These formal properties apply in all hierarchical organizations even though the content of supervision may vary considerably among organizations from collaboration with a senior author (in a research organization) to the absolute power of command in a military organization.

Formally, a hierarchical organization is a labelled set of elements (positions, organization members) and a binary relationship defined by a function sp( ) on the set of elements.

This function sp( ) and the relationship it defines must satisfy certain requirements.

Postulate Set I

Postulate 1: There exists a unique element p.

Postulate 2: For every $i \neq p$ there exists a unique j such that

(1) $j = sp(i)$ $\quad\quad i \neq p.$

Postulate 3: The relationship sp is

1) irreflexive: not $i = sp(i)$;

2) asymmetric: if $j = sp(i)$ then not $i = sp(j)$;

3) acyclic: if

$$\begin{aligned} j_1 &= sp(i) \\ j_2 &= sp(j_1) \\ &\vdots \\ j_n &= sp(j_{n-1}) \end{aligned}$$

then not $i = sp(j_n)$.

Postulate 1 states that the president is not supervised by anyone in the organization. Postulate 2 states the principle of "unified command". Every person, except the president, reports to one and only one supervisor.

Postulate 3 rules out self-supervision, mutual-supervision, and cliques as contrary to the requirements of "effective supervision" in practice.

The supervisory relationship may be represented in terms of a directed graph (digraph) such that a directed arc connects j to i whenever

$$j = sp(i).$$

Postulate Set 1 is then succinctly summarized as

Lemma: The directed graph of supervisory relationships is a tree with p as its root.

Recall that the root (or source) of a directed tree is the unique point from which a directed path exists to every other point of the tree.

This (directed) graph is usually called the organization chart. A number of properties of this digraph are discussed and proved in Supplement A. For purposes of organization theory, the following are the most interesting.

First. For every position i the organization chart contains a unique directed path from p to i. This path describes the "chain of command" that links the president to every organization member. Uniqueness means that there is exactly one "channel" to be used in communicating with supervisors and eventually the president.

Second. Consider the subset of points in the organization chart whose chain of command passes through a given element j. This subset is also a hierarchical organization (as defined on page 4) with j as its president.

Thus those parts of an organization that are headed by one person may also be studied as organizations.

Third. Three types of nodes may be distinguished which represent three types of positions in the organization:

1) a single source (or root): the president;
2) at least one sink (or endpoint): non-supervisory positions, operatives;
3) nodes having one inbound and at least one outbound arc (transit points): supervisory positions other than the president.

An alternative specification of the supervisory relationship in terms of its digraph is as follows:

Postulate Set II:

Postulate 1a:

For every node i there exists at most one predecessor node j, i.e., a j such that

$$j = sp(i);$$

Postulate 2a: The organization chart is (weakly) connected;

Postulate 3a:

The organization chart contains no (semi-) cycles.

In this characterization there is no need to introduce the distinguishing position p from the beginning. It is the result of <u>connectivity</u> and <u>acyclicity</u>.

For details cf. Supplement A, p. 19.

3. Control

The supervisory relationship may be extended by considering the set of all positions that are under the direct or indirect supervision of another position. "Direct or indirect supervision" will be called "control" and denoted by $\rightarrow$.

Definition:

An organization member i controls an organization member k if either $i = sp(k)$ "i supervises k" or there exists a chain of supervision from i to k

$$i = sp(j_1) \quad j_1 = sp(j_2), \; j_m = sp(k).$$

The relationship control is mathematically simpler than the relationship supervision.

Theorem:

The postulate set I for supervision implies the following properties of control $\rightarrow$

i. irreflexive: not $i \rightarrow i$;
ii. asymmetric: if $i \rightarrow k$, then not $k \rightarrow i$;
iii. transitive: if $i \rightarrow j$ and $j \rightarrow k$, then $i \rightarrow k$.

Proof: Since sp is irreflexive, so is $\rightarrow$.
Suppose that $i \rightarrow k$ and $k \rightarrow i$. Then there exists a cycle $i = sp(j) \; \ldots \; j_m = sp(i)$, but this is ruled out by a-

cyclicity of sp. If $i \to j$ and $j \to k$, then there exists a chain of sp from i to k. Hence $i \to k$. Q.E.D.

This theorem can be interpreted as follows. Control generates a strict pre-ordering of the organization members i. This pre-ordering may be reduced to the familiar format based on the relationship of set inclusion (see below).

With each organization member there is associated the set of members j controlled by i. This set is empty when i is not a supervisor.

Definition: Control Set S_i

$$S_i = \{j | i \to j\}.$$

Proposition: The president is the unique organization member whose control set is the entire organization (minus himself).

Proposition: The control set of any member plus this member is also an organization, i.e., a set of positions (persons) satisfying postulates I for a hierarchical organization.

The pre-ordering "control" is a simple ordering if for every pair i,k either $i \to k$ or $k \to i$. This excludes the possibility that both i and k are controlled by a third member j while neither controls the other. Therefore, the organization must be a ladder.

Definition: An organization is a ladder if all organization members can be arranged on a supervisory chain

$$p \to i_1 \to i_2 \to \ldots \to i_{N-1}.$$

Here N is the number of members of the organization. Ladders are found in pecking orders among chickens or as a "precedence ordering" in an organization (see below, Supplement B).

Corollary. When the organization is not a ladder, then there exists a pair of positions i,j such that

$$\text{neither} \quad i \to j$$

$$\text{nor} \quad j \to i.$$

As far as supervision and control are concerned nothing more than a pre-ordering is implied or needed. Any ordering of organization members by other criteria should, however, be consistent with the pre-ordering by control in order to avoid organizational conflict.

4. Rank

Some organizations deliberately avoid any ordering of members beyond that necessitated by control. Others go so far as to establish a complete strict ordering by precedence, but in such a way that precedence is consistent with and supplementary to control.

Most organizations operate between these extremes. They establish a simple ordering which, unlike precedence, is not strict. The equivalence classes in this simple ordering are called ranks.

Rank may be designated by specific titles or by classes of titles. They may be numbered. In that case economy of numbers is customary, e.g., counting by integers from zero or one upwards. Our convention will be to designate ranks r by

$$r = 0,1,2, \ldots, R.$$

where R denotes the highest rank, the rank of the president. Gaps in the numbering system will not be allowed.

The reasons for establishing a complete ordering of positions by rank are partly social and partly economic. The most important economic reason is to have a simple, consistent and transparent method for setting salaries and other compensation and economic privileges in general. The social reasons are to remove any potential conflict about any person's standing relative to the organization, in general, and to other organization members in particular.

While the pre-ordering of positions was strict, the simple ordering will be so only in the case of a ladder, when the number of ranks equals the number of positions. In all other cases there are fewer ranks than positions so that at least one rank is occupied by more than one position. This means that the simple ordering is no longer strict but contains equivalence classes

$$C_i = \{j \mid i \sim j\}.$$

The equivalence relation $\sim$ is reflexive, symmetric and transitive. Complete ordering means that

for any two organization members i,j either

$$i \rightarrow j \quad \text{or}$$

$$i \sim j \quad \text{or}$$

$$j \rightarrow i.$$

For the last statement we also write $i \leftarrow j$.
Now let rank be designated not only by title but also by numbers. Rank is a function on equivalence classes. The complete ordering of organization members is then homomorphic to the ordering of integers.

$$i \begin{array}{c} |\rightarrow| \\ \sim \\ |\leftarrow| \end{array} j \quad \text{whenever} \quad r(i) \begin{array}{c} |>| \\ = \\ |<| \end{array} r(j).$$

Consider the following rule for generating a complete ordering among organization members: Two positions i,j are considered equivalent when neither supervises the other. Is this consistent with the strict pre-ordering defined by control?

As an example consider the organization

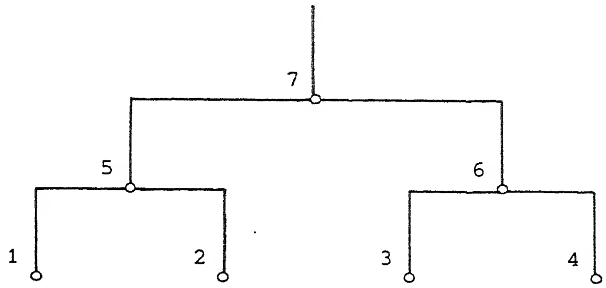


Figure 1. Organization Chart

Now

$$
\begin{array}{lll}
S_7 = \{1,2,3,4,5\} & & \\
S_6 = \{3,4\} & & S_5 = \{1,2\} \\
S_1 = \{5,3,2,1\} & & S_3 = \{5,4,2,1\} \\
 & & S_2 = \{6,1,3,4\} \\
 & & S_1 = \{6,2,3,4\}.
\end{array}
$$

Since 1 does not supervise 3 and 3 does not supervise 5 one has

$$1 \sim 3 \sim 5, \quad \text{but} \quad 5 \rightarrow 1, \quad \text{a contradiction.}$$

To obtain a ranking that is consistent with the strict pre-ordering of control $\rightarrow$ one may either count supervisory relationships or compare the sizes of control sets. We consider the second method first. Let $|S_i|$ = number of persons controlled by i and define a complete ordering of organization members by the numerical ordering of $|S_i|$. Thus

$$i \begin{matrix} |\rightarrow| \\ \sim \\ |\leftarrow| \end{matrix} j \quad \text{whenever} \quad |S_i| \begin{matrix} |>| \\ = \\ |<| \end{matrix} |S_j|.$$

Notice that the simple ordering by $>$ is consistent with the pre-ordering $\rightarrow$ defined by control, for $i \rightarrow k$ means

$k \in S_i$ which implies $S_i \quad S_k$, hence

$|S_i| > |S_k|$, hence

$i > k$.

Another natural way of setting up a complete ordering among organization members is by counting supervisory relationships--either from the top down or from the bottom up (see below) or in some intermediate way.

In a regular organization (see below Section 8) this produces the same rank as by counting the number of persons controlled. In general, the results are different and more difficult to determine when counting the number of persons controlled. Changes in employment can easily upset such a rank structure, while that based on counting supervisory links remains stable. Rank defined by the number of people controlled is, therefore, uncommon.

From now on the complete ordering of the members of an organization will be discussed in terms of its equivalence classes, the ranks. Furthermore, we assume that these ranks are numbered and that the larger number represents the higher rank.

In assigning a rank r to every position in an hierarchical organization the objective of creating comparability among all positions is achieved best by limiting the range of ranks as far as possible. The minimum range equals the longest chain of command (economy of rank numbers). Define distance

$\ell(i,j)$ = number of links in chain of command from i to j or j to i if it exists

= 0 otherwise.

Proposition: $\text{Min } R = \max_{i,j} \ell(i,j)$

$= \max_{j} \ell(p,j)$.

The range of ranks, R, depends on how many persons are supervised by each person, the span of control (see below p. 35).
In an organization with N members the largest R occurs in the case of a ladder when each person but one is a supervisor and every supervisor has one subordinate

$$R = N-1.$$

The smallest R occurs when there is but one supervisory rank. Thus

$$1 < R < N-1.$$

Better bounds can be obtained when limits are put on the number of persons supervised by any supervisor.

By convention the lowest rank is set equal to zero. The highest rank--that of president--is made as low as possible by equating it to the longest chain of command R. In assigning ranks to intermediate positions the only constraint is that for every i

$$(1) \qquad 0 \leqq r(i) \leqq r(sp(i)) - 1.$$

NB In an organization with multiple supervision the corresponding constraint is

$$0 \leqq r(i) \leqq \min_{j \varepsilon T(i)} r(j) - 1$$

where

$$T(i) = \text{set of supervisors of } i$$
$$T(i) = \{j \mid sp(i) = j\}.$$

Return to the case of unique supervision.
Now define

$$\rho(i) = \rho(sp(i)) - 1.$$

The rank assignment $\rho(i)$ is based on counting the number of links in the supervisory chain between i and the president. This is counting down. In fact,

$$\rho(i) = R - \text{number (links in chain } p \to j \to \ldots \to i).$$

Next define

$$\lambda(i) = 1 + \underset{j: \, sp(j) = i}{\underset{j}{\mathrm{Max}}} \lambda(j)$$

$$\lambda(i) = 0 \quad \text{if} \quad \{j: sp(j) = i\} = \Phi \quad \text{(the empty set).}$$

The rank assignment λ is based on assigning to nonsupervisory personnel a uniform rank of zero and counting rank from the bottom up.

Theorem:

Let $r(i)$ be any rank assignments satisfying the constraint (1).

Then $\lambda(i) \leqq r(i) \leqq \rho(i)$.

The proof is given by means of the following Lemmas:

Lemma 1: There is a longest chain of command. On this chain ranks are unique.

Proof: Otherwise R would not be equal to the length of the maximum chain.
A longest chain is also called critical. Any shorter chains is called noncritical.

Lemma 2: For all other ranks R (number of links to president) is an upper bound.

Proof: Each supervisory relationship requires at least a unit increase in rank.

Lemma 3: The number of links to the bottom end of a chain is a lower bound on rank.

Proof: The minimum rank at the end of a chain is $r = 0$. From 0 on rank must increase by at least unity for each supervisory link.

Corollary: On each noncritical chain of command the lower bound rank $\lambda(i)$ must jump by more than unity in at least one supervisory relationship.

Proof: A noncritical chain is shorter than R.

Corollary: On every noncritical chain the bottom rank under counting down is positive $\rho > 0$, hence all $\rho > 0$.

Corollary: On any noncritical chain either $r > 0$ for all positions or r jumps by more than unity between a subordinate and the supervisor.

Definition: An organizational structure is called balanced when all chains of command are critical.

Lemma 4: In a balanced structure all ranks are uniquely determined.

Proof: By definition and Lemma 1.

Example:

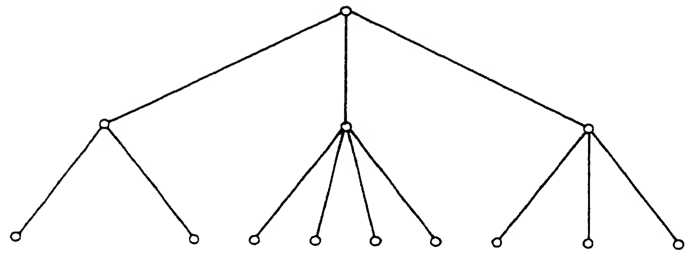


Figure 2. Balanced Organization

Corollary: In a balanced structure all nonsupervisory personnel has rank zero.

Lemma 5: Under counting up, the number of positions cannot increase with rank:

$$n_{r+1} \leqq n_r .$$

Proof: All nonsupervisory positions have rank $\lambda = 0$. For every position of rank $\lambda > 0$, the number of arrows entering from above is one and the number of arrows exiting is greater or equal to one. Hence,

$$n_{r-1} \geqq n_r , \qquad r > 0,$$

$$n_{r+1} < n_r , \qquad r = 0, \ldots, R-1.$$

Supplement A: A Graph Theoretic Model of Supervision

A natural way of describing supervision in an organization supervision is in terms of a digraph, whose points are the organization members. An arc is directed from i to j if and only if $i = sp(j)$, i supervises j. This digraph is usually called an organization chart.

The converse digraph to the organization chart describes the relationship $j = sp(i)$ "j reports to i" whose interpretation is the same. Notice that all directions are now reversed.

Postulate 2 of Section 2 states that each point i other than p has indegree one. Postulate 3 says that the organization chart contains no (directed) cycles, the digraph is acyclic.

Lemma 1: p has indegree zero.

Proof: Suppose not. Then there exists an adjacent j_1, $j_1 = sp(p)$. By postulate 2 there exists $j_2 = sp(j_1)$. Applying postulate 2 repeatedly one obtains a chain $j_n = sp(j_{n-1})$ ... $j_1 = sp(p)$ which can be continued indefinitely. Since the set of points is finite, eventually a point j_m must be repeated in the chain. This contradicts postulate 3. Hence, there is no adjacent point to p, p has indegree zero: The president has no supervisor.

Postulate 3 implies in particular
If $i = sp(j)$ then not $j = sp(i)$: irreflexiveness.
Thus arcs of opposite direction between any two points are excluded. This means

Lemma 2: the digraph is oriented.

Acyclicity implies that the points of the digraph can be pre-ordered. This is the basis of the rank system, as developed in Section 4.

Is the relationship supervision as defined by postulates 1,2,3 sufficient to hold an organization together? We show:

Theorem 3: p is a source: to every i there exists a directed path from p.

Proof: Given i there is an element j_1 adjacent to i, $j_1 = sp(i)$ by postulate 2. Either $j_1 = p$ and the proof is completed or there is an element j_2 adjacent to j_1, $j_2 = sp(j_1)$. This argument can be continued yielding a chain $j_n \ldots j_1, i$. By postulate 3 no point may be repeated in the chain. Since the point set is finite, eventually p must appear and the path from p to i is completed.

Lemma 4: p is the only source.

Proof: Let m be a second source. Then there exists a directed path from m to p contradicting Lemma 1.

Lemma 5: The organization chart is weakly connected.

Proof: Consider any two points i and j. Since p is a source there exists paths from p to i and from p to j. Together they form a semipath between i and j.

Since the digraph is oriented, it cannot be strongly connected: There cannot be a direct path from every i to every j.

Can the organization chart be unilaterally connected? Recall that a digraph is unilaterally connected if between any two points i,j there exists a path from i to j or a path from j to i.

We show:

Theorem 6: The only unilaterally connected organization chart is a ladder.
Note: A ladder is a digraph consisting of a single path.
The proof makes use of Lemma 8 below and is postponed.

Lemma 7: For every i the path from p to i is unique.

Proof: Suppose there are two paths from p to i and the points adjacent to i on these two paths are j_1 and j_2 respectively. By postulate 2, $j_1 \to i$ and $j_2 \to i$ implies $j_1 = j_2$. Continuing in this way one shows that the two paths must be identical. This lemma may be restated:

Corollary 7: There exists no semicycle between p and i. In fact, we have the stronger statement

Theorem 8: The organization chart contains no semicycles.

Proof: Suppose a semicycle exists between i and j. It cannot be a (unidirectional) cycle by postulate 3. Moving along in the path direction there must come some point k where direction is reversed. This means $m_1 = sp(k)$, $m_2 = sp(k)$ for two points m_1, m_2 adjacent to k. If $k = p$ Corollary 7 is violated, if $k \neq p$ postulate 2 is violated. Hence, no semicycle exists. Theorem 8 together with Theorem 3 imply:

Corollary 8: The organization chart is an outtree.
Recall that an outtree is a digraph with a source containing no semicycles [Harary, 3, p. 201].

Proof of Theorem 6: Let the digraph be unilaterally connected and contain at least two paths from p. Let i and j be on separate paths. A directed path from i to j cannot go through p since direction is changed at p. Hence, a semicycle exists i to j,

p to j, p to i, contradicting Theorem 8. Hence, therefore, there cannot be two paths from p: all points i,j must be on a single path from p. The digraph is a ladder.

By a well-known theorem [Harary, 3, Theorem 16.2 and 16.2', p. 200] a digraph without cycle contains at least one point of indegree zero and one point of outdegree zero. In fact, from Lemma 1 and postulate 2 one has

Lemma 1': There is one and only one point of indegree zero, p.

Points of outdegree zero will now be identified. Recall that the length of a path is the number of its arcs. Paths of maximal lengths are called maximal paths. In an organization chart maximal paths represent chains of command.

Lemma 9: A digraph of an organization with N members contains $N-1$ arcs and $N-1$ paths from p.

Proof: A tree of N points has $N-1$ arcs. By Theorem 3 and Lemma 7, to each $i \neq p$ there exists a unique path from p.

Theorem 10: All maximal paths originate in p and terminate in points of outdegree zero.

Proof: By postulate 2 any path originating in $i \neq p$ can be extended to j, $j = sp(i)$. If a termination point k has positive outdegree, the path can be extended to m, $k = sp(m)$.

Since p is a source, every point of outdegree zero is connected to p.

Corollary 10: All points of outdegree zero are endpoints of maximal paths.

Points in the organizational chart are now classified as follows:

(i) p with indegree zero and positive outdegree, origin of maximal paths.

(ii) Points with indegree one and positive outdegree: transit points of maximal paths.

(iii) Points with indegree one and outdegree zero: endpoints of maximal paths.

In an organization chart, p represents the president, transit points designate line supervisors, and endpoints denote nonsupervisors: operatives.

Transit points may be further classified as antinodes [2, p. 30] if adjacent to exactly one point and adjacent from exactly one point, and as nodes otherwise.

An antinode in an organization chart designates a line supervisor with only one subordinate. This is unusual, so that organization charts normally contain just nodes as transit points.

Remark: Each organization chart must contain points of type (i) and (iii). The only organization chart containing no transit points (ii) (line supervisors) is a star (Figure 3).

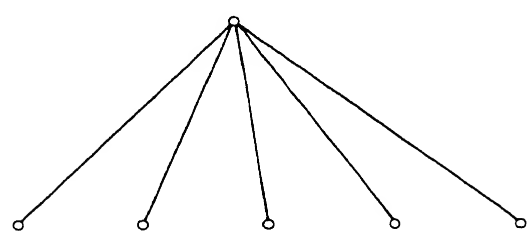


Figure 3. Star

A star contains no subgraphs other than the trivial ones: single members.

In all cases subgraphs exist that satisfy postulates 1, 2, 3 and hence constitute charts of further organizations. These may be obtained in three different ways: by deleting endpoints; as unilateral components, or as branches of the given organization chart.

A one-base is a minimal collection S of points such that each point of the digraph is either in S or adjacent from a point of S.

It is easily verified that the one-base of an organization chart is obtained by deleting its endpoints.

Lemma 11: A one-base of an organization chart is also an organization chart.

Proof: Postulates 1, 2, 3 remain satisfied.

In a similar way a two-base may be constructed by deleting endpoints of a one-base. A two-base bases exists provided after deletion of points the resulting subgraph still contains p. This process may be continued as long as the remaining organization chart contains p.

The organization theoretical interpretation is this: as set of positions consisting of the upper levels of supervisors in an organization is still an organization.

A unilateral component of a digraph is a maximal unilateral subgraph.

Lemma 12: The unilateral components of an organization chart containing point i are the paths from p through i to an endpoint.

Proof: Theorem 6.

If i is an endpoint, the unilateral component is a unique maximal path or chain of command. Paths are special organization charts, introduced above as ladders.

The branch at a point i of a digraph without semicycles may be defined as the subgraph of all points that can be reached from i. By definition a branch is an outtree.

Let $i_1, \ldots, i_m$ be the points adjacent from p. Then

Lemma 13: The organization chart is the union of p and the branches at all points adjacent to p.

Proof: Joining p to the branches at adjacent points from p makes p the source for all points.

Lemma 13 shows how the organization may be decomposed into branches such that each branch is also an outtree, an organization. Branches may be decomposed into smaller branches (divisions) and the process continued.

Lemma 14: Every transit point i is source point of a branch, and each branch is an organization.

Proof: Only at endpoints do branches consist of single points. In all other cases they are nontrivial organization charts.

Using the graph theoretical framework, other sets of postulates can be found which yield the same organizational structures. We state these postulates in terms of the digraph of an organization, the organization chart.

Recall Postulate Set II from Section 2. We restate this in graph theoretic terms.

Postulate Set II

1a. Every node i has at most one adjacent element j, $j = sp(i)$.

2a. The digraph is weakly connected.

3a. The digraph contains no semi-cycles.

Theorem 15: There exists one and only one source, i.e., an element p with no supervisor

$$i \mid i \nrightarrow p.$$

Proof: Start with any element j and consider its supervisor $i_1 = sp(j)$, the supervisor of i_1, $i_2 = sp(i_1)$, etc. This sequence must end, at the latest when all elements of the finite organization graph have been listed. For acyclicity rules out that any element may appear twice. Therefore, there is at least one source p.

Suppose there are two, p_1 and p_2. By weak connectedness a semipath exists between p_1 and p_2. The direction points away from the endpoints p_1 and p_2. At some intermediate point there must be a reversal of direction. This implies more than one supervisor for the element at which directions are reversed. Q.E.D.

This postulate set exploits the weak connectedness of the organization to relax the necessity of supervision. The distinguished element p need not be postulated here but can be proved to exist.

All other postulates sets distinguish p as a special element.

Postulate Set III

1b. The organization chart contains a distinguished point p such that no point is adjacent to p: there is no j such that $j \to p$.

2b. For every point $i \neq p$ there is a unique element j adjacent to i $j \to i$.

3b. The organization chart is weakly connected.

In this postulate set acyclicity is replaced by weak connectivity. We show

Theorem 3: p is a source.

Proof: By postulate 3 there is a semipath between p and i. At p the direction is $p = sp(j)$. Continue in this direction on the semipath until either i is reached or the direction is reversed. But reversal at m implies that $k_1 = sp(m)$, $k_2 = sp(m)$ for $k_1 \neq k_2$ contradicting postulate 2. Hence, p is

strongly connected to every i. In a similar way as before one proves

Lemma 7: To every i there is a unique path from p. Our main result is

Theorem 16: The organization chart is acyclic.

Proof: Suppose a closed cycle exists containing i. By postulate 2 any path to i must be unique. From Theorem 3 it follows that this path must contain p. Hence, the cycle must contain a point j with $j = sp(p)$ contradicting postulate 1.

The remaining properties may be proved as before.

Postulate Set III requires an organization to be held together by supervision: between any two members a connection can be found, as it turns out, via the president to whom everyone is strongly connected.

Postulate Set IV

1c. There is a distinguished point p.
2c. For every $i \neq p$ there is an adjacent point j such that $j = sp(i)$.
3c. The organization chart has no semicycles.

Notice that uniqueness is not required in postulate 2.

Lemma 1: There is no point adjacent to p.

Proof: Suppose a j_1 exists with $j_1 = sp(p)$. Then there exists $j_2 = sp(j_1)$ and so on by postulate 2. In this way the chain $j_n, j_{n-1} \cdots, j_1 \rightarrow p$ may be extended indefinitely. Since the digraph is finite, one point must be repeated eventually but this contradicts postulate 3. Hence, no j_1 exists to start this chain.

In the same way one proves

Theorem 7: To every i there is a unique path from p. This implies

Lemma 17: Each $i \neq p$ has a unique adjacent point j, $j = sp(i)$. The remaining properties follow as before.

Postulate set 3 shows that uniqueness of supervision is implied when acyclicity is strengthened to mean exclusion of semicycles.

Postulate Set V

Consider an organization with N members and an organization chart defined as follows:

1d. p is the only point with indegree zero.

2d. The digraph is weakly connected.

3d. It contains $N-1$ arcs.

Theorem 8: The organization chart contains no semicycles.

Proof: A (weakly) connected graph with semicycles contains at least N arcs.

With Theorem 8, all postulates in set IV are satisfied.

This postulate set achieves uniqueness of supervision by limiting the number of supervisory relationships.

In conclusion we remark that the number O_N of different organization charts with N members when all members are distinguished is equal to the number of labelled trees [Harary, 3, p. 179]

$$O_N = N^{N-2}.$$

The number of different organization charts with only the president identified is the number of unlabelled rooted trees (with root p). It is usually stated in terms of its counting series

$$T(x) = \sum_{N=1}^{\infty} T_N x^N$$

where T_N is the number of rooted trees with N elements. In principle, T_N may be found from Cayley's functional equation

$$T(x) = x \exp \sum_{n=1}^{\infty} T(x^n).$$

The following numbers are taken from [Harary, 3, p. 232]

N	3	4	5	6	7	8	9	10	11	12	15
T_N	2	4	9	20	48	115	286	719	1,842	4,766	87,814

N	20	25
T_N	12,826,228	2,067,174,645

Supplement B: Precedence

For some purposes ranking in terms of equivalence classes under supervision (by counting the number of supervisory relationships) is not enough. Within each rank a further ordering is required resulting in a ladder, i.e., a complete strict ordering. Examples

1. Order of succession in the government (president, vice president, senate majority leader, speaker of the house, secretary of state); also in corporations.
2. In the military: ranking officer (rank and seniority). German universities: rank and date of appointment to rank.

Precedence is partly honorific: who is the acting, i.e., ranking representative and recipient of honors accorded to the organization. It is operational in regard to succession to the presidency. Precedence supplements rank, it can never be contradictory to rank. It may be based on supplementary criteria such as

(i) the actual number of persons under a position's control the size of the control set;
(ii) seniority;
(iii) merit, i.e., achievement;
(iv) importance of area or department.

Rank and precedence even when earned by achievement, are basically descriptive: once obtained they are held regardless of achievement.

Rank Compared

To degree of kinship. In North European cultures (and under Roman law) the number of parent-child relationships is counted, to determine the degree of kinship.

Relative	Degree of Kinship
Father, Mother	1
Son, Daughter	1
Brother, Sister	2
Grandfather, Grandchild	2
Uncle, Aunt	3
Nephew, Niece	3
Cousin	4

In a Royal house, rank is defined as degree of kinship to the sovereign.

Supplement C: Communication Aspects of Rank

In small organizations every member can have access to every other member without undue interference. However, in large organizations communications must go through "channels" for three basic reasons.

i) It may not be known who is the right person to be contacted.

ii) It is more efficient to bundle messages.

iii) There would be too many messages.

The supervisory structure offers a natural communications network through which communications can be channelled and filtered so that the load of incoming messages is held at an acceptable level. In practice, the official channels are always supplemented by informal links through personal friendship by which the official path can be short circuited. Compared to the number of possible pairwise links, these unofficial links are not numerically important at the lowest level, but can achieve effective short cuts to path length for messages between members of higher ranks.

The following propositions are based on the supervisory structure as the official communications network.

Suppose that all direct communications are between supervisors and their immediate subordinates. Communications paths are obtained by stringing together such links between pairs of supervisor and subordinate.

Proposition 1: Between any two organization members i,j there exists a unique communication path such that each link in the path is traversed once and only once.

Proof: Take the unique chain of command from i to the president and from j to the president and eliminate those portions that have been traversed twice.

N.B.: With informal links added to the communications network, there may be additional informal paths of shorter length.

Definition: The distance between two organization members i,j is the number of links on their unique communications path.
N.B.: In the union of the formal and informal communication networks distance is the minimal length for all existing communication paths.

Corollary: The distance between an organization member of rank r and an organization member of rank ρ is not larger than $2R - r - \rho$.

Definition: The height R of an organization is the length of the longest chain of command.

Proposition: In an organization of height R the diameter or largest distance is $\leqq 2R$, and = if there are at least two chains of command of length R.
N.B.: In a balanced organization the diameter is $2R$.

Definition: The eccentricity of a point i in the organization is the largest distance from i to any member of the organization.

Proposition 3: In a balanced organization the eccentricity of any position of rank R is $2R - r$.

Definition: A point of minimal eccentricity is called a center.

Proposition 4: A tree of even diameter has a unique center, a tree of odd diameter has two adjacent centers.

Proposition 5: In an organization with at least two chains of command of equal maximal length R the president occupies the unique center.

Corollary: In a balanced organization the president occupies the center, i.e., has smallest maximal distance to any organization member, equal to R.

Remark: The president has the smallest average distance to all organization members if for every given rank each supervisor has the same number of subordinates.

In conclusion we remark, that an organization chart may designate three different types of relationship

supervision and control

official channels of communication

career tracks.

When these relationships are isomorphic, they may be represented by the same graph.

5. <u>Span of Control</u>

5.1 Definition. The relationship "supervision" is not one-to-one. By Postulate 2 of Section 2, each organization member, except the president, has one and only one supervisor. Supervision would give rise to unnecessary chains of command, if a supervisor had only one subordinate. The number of immediate subordinates of a full-time supervisor is called his/her span of control s. From now on this span is assumed to be at least 2,

$$s \geqq 2.$$

Recommended values are given in the "management" literature [e.g., Koontz-O'Donnell, 1959]

$$3 \leqq s < 8 \qquad \text{for supervising managers}$$

$$8 < s < 15 \qquad \text{for supervising operatives.}$$

The question of what determines the span of control and what spans are optimal is discussed in Chapter IV.

Without a bound on the span of control, organizations could be made simple by appointing one member supervisor (the president) and all others to operative positions.

In the graph representation of an organization--the organization chart--each node has a degree defined as the number of

edges incident to this node. The degree is one for operatives (located at endpoints); it equals the span of control for the president (the root). The span of control is the number of positions directly supervised by a supervisor. The degree equals one plus the span of control for nonpresidential supervisors (transit points).

Thus, if the span of control s is equal for all supervisory positions, the node representing the president may be identified as the unique node of degree $s+1$.

The graph representation of the supervisory (or control) structure of organization suggests that two types of imbalance may occur:

unequal length of chains of command and

uneven spans of control.

Other things equal a design is improved when a larger span of control is reduced by shifting a position so as to increase a smaller span of control. This is particularly apparent when the maximum span of control is thereby reduced. An organization is called balanced when all its chains of command have equal length. This implies

1) The presidential position is the unique center of the organizational graph.
2) Rank assignments under counting up and counting down are equal.

5.2 Constant Span of Control

In the remainder of this section we consider organizations with constant span of control s. This is interpreted to mean that any organization member can supervise at most s other organization members. If he/she supervises

$$m < s$$

members, then he/she can perform an amount $\frac{s-m}{s}$ of operative work. (All work is measured in person units, assumed to be equal for all organization members.) Such a person would do both supervision and operative work.

To begin with we shall allow part-time employment as well. The total amount of work any particular person can be hired to do is then $\leqq 1$.

Let the task of the organization be given and consist of Q person units of operative work. We ask:

What is the minimum amount of total supervisory work required by Q units of operative work?

What is the minimum size of the organization and how does it depend on the organizational design, i.e., the way in which operative and supervisory work are allocated and supervisory relationships are arranged?

Our principal result is the following:

<u>Theorem 1</u>

Let the span of control s and the size of the task Q be given. Then the required amount of supervision M is uniquely determined and independent of organizational design

$$M = \frac{Q-1}{s-1} \tag{1}$$

<u>Corollary 2</u>

The size of the organization is determined by

$$N = \frac{sQ-1}{s-1}$$

<u>Proof:</u> The total number of organization members is $M + Q$, but only $M + Q - 1$ require supervision. The number of supervisors required (including the president) is therefore

$$M = \frac{1}{s}(M + Q - 1) \tag{2}$$

Solving (2) for M yields

$$M = \frac{Q-1}{s-1} \qquad (1)$$

The number of supervisors is thus a linear function of the number of operatives Q. For large organizations $1 \ll Q$ so that approximately

$$M \doteq \frac{Q}{s-1} \qquad (3)$$

Thus each operative gives rise to an administrative force totalling

$$\frac{1}{s} + \frac{1}{s^2} + \frac{1}{s^3} + \ldots = \frac{1}{s-1}.$$

The fact that the president needs no supervision actually terminates this series after at most R steps, and this results in the corrective term $\frac{-1/Q}{s-1}$ in (1) omitted from (3). From (1), the size of the organization is obtained by considering

$$N = Q + M$$

$$N = \frac{sQ-1}{s-1} \qquad (4)$$

using (1).

For a large organization we may approximate (4) by

$$N \doteq \frac{Q}{1 - \frac{1}{s}}. \qquad (5)$$

The factor $\frac{1}{1 - \frac{1}{s}}$ is sometimes called the <u>organizational multiplier</u>. It shows the total employment generated in an organization by one unit of outside work Q. The organizational multiplier is sensitive to s when s is small. Thus using a span

of control $s = 3$ when $s = 4$ would have been adequate increases the size of the organization by a factor

$$\frac{1}{1 - \frac{1}{3}} - \frac{1}{1 - \frac{1}{4}} = \frac{3}{2} - \frac{4}{3} = \frac{1}{6}$$

The fact that the total requirement of supervision should be independent of organizational design is surprising. Compare, however, the following possibilities of organizing supervision for $Q = 3$ when $s = 3$. The amount of operative work done by each organization member is listed in parentheses.

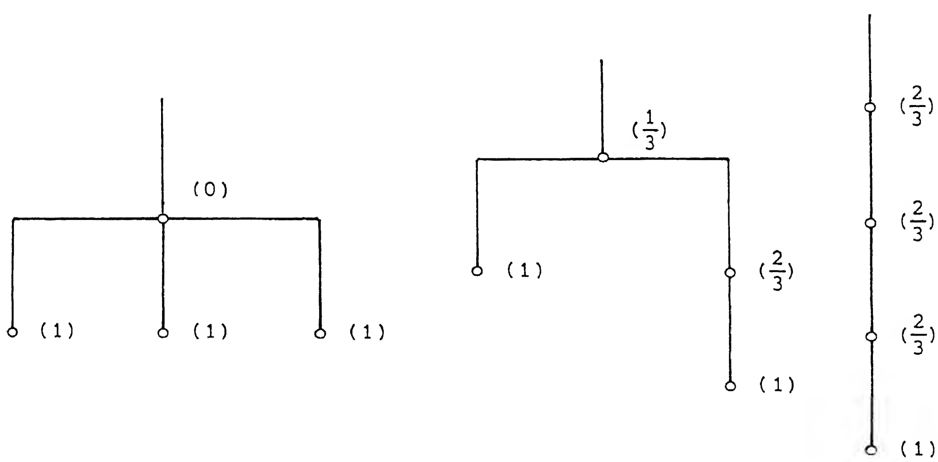


Figure 4. Organizational Designs for $Q = 3$, $s = 3$

So far, fractional (i.e., part-time) employment was admitted as well as a splitting of time between operative and supervisory roles. In this section we examine the implications first of full-time employment and secondly of full-time assignments to supervisory and operative work.

As before, Q and s are assumed given. In view of (1) and (4) full-time employment is possible, if and only if

$$Q \text{ is integer and } s-1 \mid Q-1 \qquad (6)$$

It is easy to show that (1) also implies that the solution of (4) is integer.

Can the integer requirements of supervision always be implemented by full-time assignments of personnel to supervisory and operative positions? The answer is positive and how it is done is shown in Section 7.

When the integer or divisibility requirement (6) is not satisfied, then either Q or M is not an integer; however, N is integer provided

$$s-1 \mid sQ-1 \qquad (7)$$

Since s is always an integer, condition (7) requires that sQ be integer.

When (7) is violated, full-time hiring implies an organization of size

$$N = \left| \frac{sQ-1}{s-1} \right| \qquad (8)$$

where

$$\{x\} = \text{smallest integer} \geqq x.$$

If the number of operatives Q is integer, then (8) implies a number of full-time supervisors equal to

$$M = \left| \frac{Q-1}{s-1} \right| \qquad (9)$$

6. Estimation

The actual span of control can always be read off an organization chart. Suppose, however, only the numbers of positions of various ranks n_r are given. In this case a lower bound on the average span of control may be inferred. If more information is available, e.g., that the organization is balanced, more precise values may be calculated.

Suppose first that all supervisors and all operatives can be identified. Then M and Q are known. From (5.3)

$$(s-1)M = Q-1 = N-M-1$$

$$s = \frac{N-1}{M} \tag{1}$$

(For an alternative derivation cf. Supplement E.) This is true, for instance, when all persons of rank r^* or lower are operatives and all persons of rank above r^* are supervisors. In this case

$$M = N_{r^*+1}$$

Here N_r is defined as $N_r = \sum_{i=r}^{R} m_i$.

We consider next the case where only a set containing the supervisors is known.

Theorem: Let G be a subset of members of an organization and let H contain the set M of all supervisors of G. Then

$$s \geqq \frac{|G| - 1}{|H|} \tag{2}$$

<u>Proof</u>:

$$|M| = \frac{1}{s}(|G| - 1)$$

if G contains the president and the members of G supervise no persons not belonging to M. Otherwise,

$$|M| > \frac{1}{s}(|G| - 1).$$

Combining the inequality and equation yields

$$s \geqq \frac{|G| - 1}{|M|} \geqq \frac{|G| - 1}{|H|} \quad \text{since} \quad H \quad M.$$

<u>Application</u>: Let G = set of persons of rank $\geqq r$

$$|M| = N_r = n_r + n_{r+1} + \dots + n_{R-1} + n_R$$

H = set of persons of rank $\geqq r+1$

$$|H| = N_{r+1}.$$

Then

$$s(r) \geqq \frac{N_r - 1}{N_{r+1}} \tag{3}$$

where $s(r)$ denotes the average span of control of supervisors above rank r.

The average span of control for the entire organization is then given by

Corollary:

$$s \geqq \operatorname*{Max}_{r} \frac{N_r - 1}{N_{r+1}} \tag{4}$$

Example: Library of Congress [The Budget, 1982]

n_r	N_r	s(r)
1		
	1	
1	2	1
1	3	1
7	10	3
3	13	1.2
9	22	1.6
45	67	3
57	124	1.8
188	312	2.5
218	530	1.7
322	852	1.6
616	1468	1.7

The estimated span of control at levels GS 18 and higher is $s = 3$.

Suppose it is known that all positions above a certain rank r^* are full-time supervisors and that the span of control is constant. Then the estimate (3) is exact and r^* is determined as the maximizer in (3).

Example: Office of Management and Budget [The Budget, 1982]

r	n_r	N_r	$\hat{s}(r)$
Exec II	1	1	
Exec III	1	2	2
ES 6	9	11	5
5	6	17	1.46
4	65	82	4.76
3	4	86	1.03

The estimated span of control is 5. It applies at the two upper levels. The second estimate 4.76 is close enough to permit the inference, that $r^* =$ ES 4 so that all personnel above this rank is purely supervisory.

In a <u>balanced organization</u>, the supervisors of personnel in rank r are all found in rank $r+1$. Therefore, the average span of control for supervisors of rank r is

$$s_r = \frac{n_{r-1}}{n_r} \qquad r = 1, \ldots, R. \tag{7}$$

The average span of control for the entire organization may be estimated as the geometric mean of (7).

$$\hat{s} = \sqrt[R]{s_1 \cdot s_2 \ldots s_R}$$

$$= \sqrt[R]{\frac{n_o}{n_1} \cdot \frac{n_1}{n_2} \cdot \ldots \frac{n_{R-1}}{1}}$$

$$\hat{s} = \sqrt[R]{n_o} \tag{8}$$

In a nonbalanced organization not all operative positions are in rank zero. The span of control is then underestimated by (8).

$$s \geqq \sqrt[R]{n_o} \tag{9}$$

A sharper bound for s can be found by eliminating from H all non-supervisory personnel. Let x_r denote supervisors of rank r and

$$X_r = x_r + x_{r+1} + \ldots + x_R.$$

Then

$$s \geqq \frac{N_r - 1}{X_{r+1}}. \tag{5}$$

In some cases all ranks $r = 0, 1, \ldots, r_o$ are occupied by non-supervisory personnel. Then

$$s \geqq \frac{N_o - 1}{N_{r_o+1}}. \tag{6}$$

which is a weaker estimate than (1).

7. Assignment

Can the supervisory requirements always be implemented without organizational slack? In particular, when

$$s-1 \mid Q-1 \qquad (1)$$

how can one design an organizational structure composed of full-time operatives and full-time managers as supervisors?

There are, in fact, many designs which implement (5.3), one is described in Supplement G. In fact, there are as many as there are trees with the properties mentioned above: having Q endpoints, one internal point of degree s and all other internal points of degree $s+1$.

To see this we recall that for any graph

$$\text{numbers of arcs} = \sum_{\text{nodes}} \text{degrees.}$$

Recall also that in a tree the number of arcs is one less than the number of nodes. Thus

$$2(N-1) = Q + (s+1)(M-1) + s$$

$$= Q + (s+1)(N-Q-1) + s$$

yielding the correct size of the organization, viz.,

$$N = \frac{sQ-1}{s-1}. \qquad (3)$$

which is identical with (5.2).

Among all designs satisfying the requirements (5.2) one can, in fact, be found such that the ranks assigned to supervisors have the counting up property of being minimal.

The algorithms for this is as follows. All Q operatives are given rank zero and a maximum number of supervisors is assigned to these and given rank one. The number of supervisors of rank one is then equal to

$$[\frac{Q}{s}] = n_1 .$$

This leaves $Q - s\,[\frac{Q}{s}] = k_1$

positions unsupervised.

Next, we determine a maximum number of supervisors of rank 2

$$\left|\frac{n_1 + k_1}{s}\right| = n_2$$

leaving

$$n_1 + k_1 - s\left|\frac{n_1 + k_1}{s}\right| = k_2$$

positions unsupervised.

In general,

$$n_{r+1} = \left|\frac{n_r + k_r}{s}\right| \tag{4}$$

$$k_{r+1} = n_r + k_r - sn_{r+1} . \tag{5}$$

Since the numbers of positions falls by at most $\frac{1}{s}$ as one moves up, it follows that

$$Q \leqq s^R \qquad \text{or}$$

$$R \geqq \frac{\log Q}{\log s}. \tag{6}$$

This can be continued until $n_R = 1$.

It is easy to show that $n_r + k_r$ and hence n_{r+1} decrease with r as long as $n_{r+1} > 1$.

Moreover, one shows by induction that

$$s-1 | (n_r + k_r - 1)$$

for all positive values of the parenthesis. From these two facts it follows that eventually

$$n_R = 1 \quad \text{and} \quad k_R = 0.$$

It is clear that this algorithm produces a rank assignment by "counting-up" and that the resulting numbers of positions at each level and thus the design are unique.

Supplement D
Integers and Slack

A minor source of inefficiency occurs when the divisibility requirement in (12) is violated. It is then also violated in (13), and vice versa. The resulting "organizational slack" or excess capacity of supervisors can never exceed unity.

The algorithm described for assigning supervisors in such a way that these supervisors have minimal rank may then fail to produce an efficient organization by ending with a presidential rank that is too high. One must, therefore, reconsider, whether organizational slack should be introduced at a lower level, eliminating a k_r and thereby effectively reducing the top rank required. The reader may verify this in the example of Figure 2.

Organizational slack is also introduced when a balanced organization of minimum size is sought that is needed to supervise a given work force of Q operatives.

The maximal slack is incurred for task sizes

$$Q = s^{R-1} + 1 \tag{1}$$

In this case increasing the task from $Q-1$ to Q requires no fewer than R additional supervisors. The slack

$$R = \frac{\log (Q-1)}{\log s} + 1 \tag{2}$$

can be large for a small organization. Thus, when $s = 2$ and $Q = 3$, increasing the task from 2 to 3 requires 3 additional

positions: a 50% increase in work gives rise to a 100% increase in personnel.

But the maximal ratio of slack to output in case (1)

$$R \cdot s^{-R} \tag{3}$$

decrease rapidly with R.

<u>Lemma 1</u>: Any slack can be assigned to a single organization member.

<u>Proof</u>: See algorithm of Section 7 for minimizing ranks.
N.B.: It is not always possible to assign this slack to the president and have minimal R as the example of Figure 5 shown.

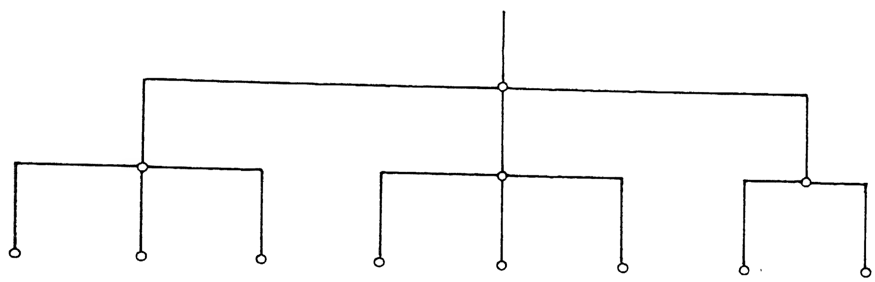


Figure 5. Organizational Slack

<u>Lemma 2</u>: The amount of slack, i.e., unused capacity in an organization using $\{\frac{Q-1}{s-1}\}$ supervisors is

$$\frac{s-1}{s} \cdot \left(\left\{\frac{Q-1}{s-1}\right\} - \frac{Q-1}{s-1}\right).$$

Proof:

Available as supervisors M

needed supervision $\frac{1}{s}(M+Q-1)$

$$\text{slack} = M - \frac{1}{s}(M+Q-1)$$

$$= \frac{s-1}{s} M - \frac{Q-1}{s}$$

$$= \frac{s-1}{s}\left(\left\{\frac{Q-1}{s-1}\right\} - \frac{Q-1}{s-1}\right) \qquad \text{Q.E.D.}$$

The slack can be distributed in units of $\frac{1}{s}$ among certain (but not all) supervisors. It can also be concentrated in one position, but not always in the president.

The correction factor $\frac{s-1}{s}$ is due to the fact that increasing the supervisors from $\frac{Q-1}{s-1}$ to $\left\{\frac{Q-1}{s-1}\right\}$ also raises the amount of supervision now required and thereby reduces the slack.

Corollary 3: The slack if positive is bounded by

$$\frac{1}{s} < \text{slack} \leqq \frac{s-2}{s}.$$

Corollary 4: The ratio of supervisors to operatives is bounded by

$$\frac{1}{s-1} - \frac{1}{(s-1)Q} < \frac{M}{Q} \leqq \frac{1}{s-1} + \frac{(s-2)^2 - 2}{s(s-1)Q}$$

These bounds rapidly approach the value

$$\frac{M}{Q} = \frac{1}{s-1}.$$

If the slack is at the presidential level, then $R > 1$ implies $Q = ms+i \quad m > 1$ and slack $\frac{m+i}{s}$.

Now

$$Q-1 = (m-1)s + s-1 + i$$

$$Q = ms+i$$

and

$$\frac{N}{Q} = \frac{1}{1 - \frac{1}{s}} Q.$$

Corollary 5: For $s = 2$ there is never any slack.

Supplement E

Average Span of Control and More Graph Theory

The supervisory structure of an organization was discussed in terms of a digraph. For certain purposes the directions of arcs may be ignored so that an ordinary (undirected) graph is obtained.

Notice, first, that this graph of an organization is a tree. The N points of the tree are joined by $N-1$ supervisory links and by $\frac{N(N-1)}{2}$ (undirected) unique paths. There are at least two and at most $N-1$ endpoints, i.e., points of degree one $(N > 2)$. The first bound is reached by a ladder, the second by a star.

Let $I = \{i \mid i \neq p,\ \deg(i) > i\}$

be the set of points representing line supervisors. The span of control of any $i \varepsilon I$ is defined as

$$s_i = \deg(i) - 1 \qquad (1)$$

$\deg(i)$ is the degree, the number of lines incident with i. For p the span of control is defined as

$$s_p = \deg(p). \qquad (2)$$

Let n_o be the number of endpoints. Then

$$\bar{s} = \frac{s_p + \sum_{i \varepsilon I} s_i}{N-n_o} \tag{3}$$

is the average span of control. We show that $\bar{s}$ depends only on N and n_o.

A well-known formula of graph theory states [Harary, 1971, p. 14].

$$\sum \text{degrees} = 2 \cdot \text{number of lines} \tag{4}$$

Substituting

$$s_p + \sum_{i \varepsilon I} (s_i+1) + n_o = 2(N-1)$$

$$s_p + \sum_{i \varepsilon I} s_i + N-n_o - 1 + n_o = 2(N-1)$$

$$s_p + \sum_{i \varepsilon I} s_i = N-1. \tag{5}$$

The sum of all spans of control equals the number of members to be supervised $N-1$.

Furthermore, by (3)

$$\bar{s} = \frac{N-1}{N-n_o}. \tag{6}$$

<u>Theorem 1</u>

The average span of control is independent of structural detail. In particular, it is independent of the location of p.

Consider next, the distance $d(i,j)$ between members of an organization and the excentricity of a point defined as

$$e(i) = \operatorname*{Max}_{j} \ d(i,j). \qquad (7)$$

Excentricity assumes its minimum at the center. The center of a tree consists of either one point or two adjacent points [Haravy, 1971, Theorem 4.2, p. 35].

We ask: Under what conditions is the element p identified by the structure of the organization's graph?

Distinguished points in a tree are either endpoints or a unique central point.

Now an endpoint having degree one would imply a presidential span of control of one. In effect, all members of the organization would report to a subordinate of the president, a degenerate case that will be ignored.

In the following p is assumed to be the unique central point of the organization's tree.

Lemma 2: p is a unique central point if and only if there exist at least two maximal paths from p, i.e., paths to endpoints which have equal maximal length.

Proof: If there are two central points, the successive deletion of endpoints leaves only one maximal path (of length one). Hence, more than one maximal path is necessary. It is sufficient for successive deletion of endpoints reduces the graph to just one point when more than one maximal path are present.

The excentricity of the center is the radius of the graph. The lemma implies

Corollary 3: The diameter of the organization's graph equals twice its radius.

Thus, for the president to be identified on an undirected organizational graph, there must exist at least two longest chains of command. By definition of a center, which the president occupies, the president is closest to the remainder of the organization in the sense that the length of the longest communications chain is minimized.

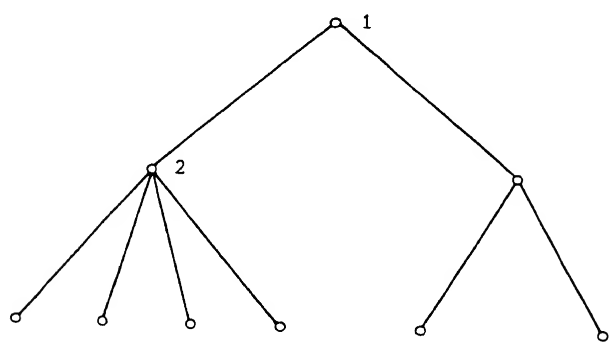
An organization in which all graphs from center to endpoints have equal length is called balanced. Its chains of command are of equal length.

As an alternative measure of closeness to members of the organization, the total or the average distance to organization members may be considered.

$$D(i) = \sum_j d(i,j) \quad \text{and} \tag{8}$$

$$\bar{d}(i) = \frac{1}{N-1} \sum_j d(i,j). \tag{9}$$

However, these are not necessarily minimized at a center, as the example of Figure 2 shows.



D(1) = 14
D(2) = 13

Figure 6 .
Distances from Centre are not minimal

However, in a balanced organization with constant span of control in each rank, total and average distance are minimized at the center.

Let R be the organization's radius, and s the uniform span of control. Then

$$D(p) = \frac{s}{(s-1)^2} [Rs^{R+1} - (R+1)s^R + 1] \qquad (10)$$

$$\bar{d}(p) = \frac{Rs^{R+1} - (R+1)s^R + 1}{(s-1)(s^R-1)} \qquad (11)$$

and these values are the minimal ones for all points in the graph. $\bar{d}(p)$ is, in fact, independent of the presidential span of control.

Notice that (10) is an integer but (11) may be fractional.

For large R or s, $\bar{d}$ is approximated by

$$\bar{d}(p) = R - \frac{1}{s-1} \qquad (12)$$

which shows that the minimum average distance is close to the radius, the maximal distance from the center, and closer the larger is s.

Supplement F. Support Structure

We consider an organization, like the Army, which contains a support structure supplying necessary services to both supervisors and operatives.

Let each person in the organization require the services of $\frac{1}{m}$ support persons. As before let

Q = number of operatives

M = number of supervisors

B = number of support personnel

N = size of the organization.

By hypothesis

$$M = \frac{1}{s}(Q + M+B-1) \qquad (1)$$

$$B = \frac{1}{m}(Q+M+B) \qquad (2)$$

A straightforward calculation yields

$$B = \frac{Q+M}{m-1}$$

$$N = Q+M+B = \frac{Q+M}{1-\frac{1}{m}}$$

$$M = \frac{1}{s}(N-1) = \frac{1}{s}\left(\frac{Q+M}{1-\frac{1}{m}} - 1\right)$$

$$M = \frac{m(Q-1) + 1}{(s-1)(m-1) - 1} \qquad (3)$$

$$B = \frac{sQ-1}{(s-1)(m-1) - 1} \tag{4}$$

$$N = \frac{sQ-1}{s(1 - \frac{1}{m}) - 1} \tag{5}$$

As an illustration suppose that

$s = 3 \qquad m = 2.$ Then

$$M = 2Q-1$$

$$B = 3Q-1$$

$$N = 6Q-2.$$

The combination of support and supervision can thus blow up the size of an organization considerably.

Supplement G. Maximizing Average Rank

Under the constraint that the maximal chain contains no rank jumps. Let

$$Q-1 = m \cdot (s-1)$$

$$Q = s + (m-1)(s-1).$$

This can be implemented by assignment the work load Q to workers of rank r, q_r as follows.

$$q_o = s$$

$$q_r = s-1 \qquad r = 1, \ldots, R-1$$

$$x_r = 1 \qquad r = 1, \ldots, R.$$

The resulting rank occupancies

$$n_r = q_r + x_r$$

are then as follows

$$n_r = \begin{vmatrix} 1 & \quad r = R \\ s & \quad r = 0, \ldots, R-1. \end{vmatrix}$$

Having the minimal number of one supervisor in each rank, hence each rank below R occupied by no more than s members maximizes R. R would be unchanged when all operatives are downgraded to $r = 0$. This arrangement assigns the smallest numbers q_o to the lowest rank, the smallest to the next rank, etc., implying a maximal number in the higher ranks. Moreover, R was maximal itself:

By having no more than one supervisor per rank (the minimal number) the number of ranks is maximized.

Proof: (induction with respect to m).
Notice first that rank jumps are inefficient, i.e., will lower ranks when occurring outside a maximal chain and are not permitted on a longest chain. For $m = 1$, $s = q$ and the only assignment is $n_1 = 1$, $n_o = q_o = s$. Now consider any $m > 1$. The minimum number of personnel at rank zero is s. This requires one supervisor at the next level. The work load has been reduced by s but one supervisor has been added. The resulting organization, therefore, has a work load

$$Q = 1 + (m-1)(s-1).$$

By induction hypothesis this organization consists of layers of s persons. Together with the bottom stratum of s persons this makes up an organization as asserted, i.e., an s-ladder. Q.E.D.

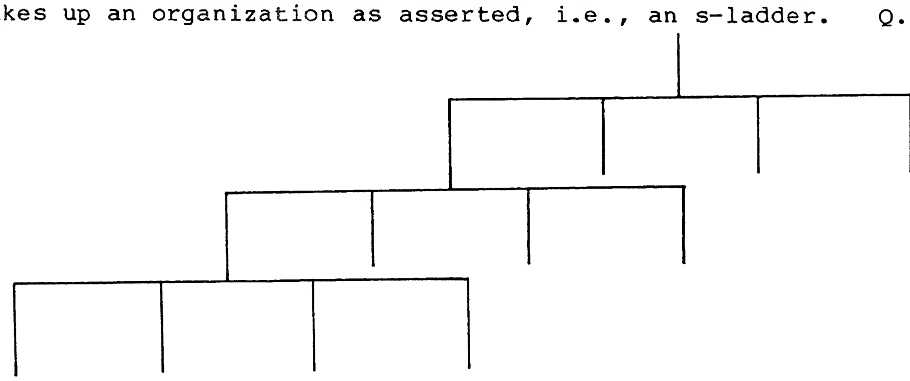


Figure 7. s-ladder

Problems

1. Let s_r be the (average) span of control for supervisors of rank r. Show that counting up in the assignment of rank implies

$$\frac{n_{r-1}}{n_r} \leqq s_r$$

and that counting down implies

$$\frac{n_{r-1}}{n_r} = s_r .$$

2. Suppose that the span of control for supervising supervisors is s and for supervising operatives is s_o. Show that the amount of supervision generated by Q operatives is

$$M = \frac{sQ-s_o}{(s-1)s_o}$$

and the size of the organization required is

$$N = \frac{(s + \frac{s}{s_o} - 1)Q-1}{s-1} .$$

(Note: It is assumed that the president does no operative work.)

3. Suppose that it is considered desirable that first line supervisors also do some operative work and that managers, in general, do some (supervisory) work that is assigned to managers under their (immediate) supervision. What does this imply for

the size of the organization

the number of operatives

the number of supervisors

the number of ranks required?

III. Costs and Scale

8. Regular Organization

A balanced organization with constant span of control is called regular.
In particular, an organization is called regular when all its operatives have rank zero and all supervisors have a constant span of control s.

An organization of this type has an ideal structure which may be characterized as follows.
It has no slack and makes maximal use of the capacity to supervise: No person of rank $r > 0$ performs operative work.
Therefore:

i) For a given s and R the size of the task Q is maximized.

ii) For a given s and Q the maximal rank R is minimized.

iii) For a given Q and R the maximal span of control s is minimized.

To prove property i) observe that

$$n_R = 1$$

$$n_{R-1} \leqq s \quad \text{and} \quad = \quad \text{for regular organization}$$

$$n_{R-2} \leqq s^2 \quad \text{and} \quad = \quad \text{for regular organization}$$

$$\vdots$$

$$n_r \leqq s^{R-r}$$

$$\vdots$$

$$n_o \leqq s^R \quad \text{and} \quad = \quad \text{for regular organization}$$

Properties ii) and iii) are corollaries.

The regular organization is, in fact, characterized by

$$n_r = s^{R-r} \qquad r = 0,1,\ \dots,\ R \tag{1}$$

The image of a regular organization as a pyramid is, therefore, misleading. The basis is much broader, since the number of positions increases exponentially as one moves down the hierarchy of ranks.

The number of persons of rank r or higher rank is

$$N_r = 1 + s + \dots + s^{R-r} = \frac{s^{R-r+1} - 1}{s-1} \tag{2}$$

Consider average rank

$$\bar{r} = \frac{\sum_{r=0}^{R} rn_r}{\sum_{s=0}^{R}} = \frac{1}{s-1} - \frac{R+1}{s^{R+1}-1} \tag{3}$$

Average rank increases with maximal rank R, but remains bounded and approaches rapidly this bound

$$\bar{r} \leqq \frac{1}{s-1} \leqq 1 \quad \text{for} \quad s > 1 \tag{4}$$

This bound $\frac{1}{s-1}$ is itself not larger than unity when the span of control is at least 2.

That the average rank of a regular organization is minimal among all organizations with given Q and s may now be shown as follows.

We begin by deriving a lower bound on average rank in terms of s, R and N.

Supervisory requirements in a ranked organization imply, cf. (6.3),

$$N_{r+1} \geqq \frac{1}{s}(N_r-1) \tag{5}$$

Subtract this inequality from $N_r = N_r$ to obtain

$$N_r - N_{r+1} \leqq \frac{s-1}{s} N_r + \frac{1}{s}$$

Summation over r yields

$$\sum_{r=0}^{R} (N_r-N_{r+1}) \leqq \frac{s-1}{s} \sum_{r=0}^{R} N_r + \frac{R+1}{s}$$

Observe that the left-hand side equals

$$\sum_{r=0}^{R} (N_r-N_{r+1}) = \sum_{r=0}^{R} n_r = N \tag{6}$$

and that

$$\sum_{r=0}^{R} N_r = \sum_{r=0}^{R} (r+1)n_r = N(\bar{r}+1) \tag{7}$$

Substituting (6) and (7) in (5) yields

$$N \leqq \frac{s-1}{s} N + \frac{s-1}{s} N\bar{r} + \frac{R+1}{s} \qquad \text{or}$$

$$\bar{r} \geqq \frac{1}{s-1} - \frac{R+1}{(s-1)N} \tag{3a}$$

As we have seen, this inequality is binding for a

regular organization with $N = \frac{s^{R+1}-1}{s-1}$

$$\bar{r} = \frac{1}{s-1} - \frac{R+1}{s^{R+1}-1} \tag{3}$$

Suppose we wish to minimize average rank $\bar{r}$ for given N.

Now $\quad \bar{r} = \frac{1}{N} \sum_{r=1}^{R} N_r \quad$ where R may be arbitrarily large.

The minimization problem is then

$$\text{Min } \frac{1}{N} \sum_{r=1}^{R} N_r$$

subject to $\quad N_{r+1} \geqq \frac{1}{s} (N_r - 1) \qquad r = 0,1,... \qquad (4)$

Since $N_0 = N$ is given, this minimum is obtained when the "=" sign holds in all constraints (4).

$$N_{r+1} = \frac{1}{s} (N_r - 1) \tag{7}$$

Equation (7) is solved by

$$N_1 = \frac{1}{s} N - \frac{1}{s} \quad \text{etc.,}$$

or, beginning with $r = R$

$$N_R = 1$$

$$N_{R-1} = 1 + s\,N_R = 1 + s$$

$$N_{R-2} = 1 + s + s^2$$

from which

$$n_r = N_r - N_{r+1} = s^{R-r} \qquad (1)$$

This proves

<u>Lemma</u>: A regular organization has smallest average rank among all organizations of given size N.

Consider next the average distance between the president and all members of the organization.

$$\bar{d}_R = R - \bar{r}$$

Substituting for $\bar{r}$

$$\bar{d}_R = R - \frac{1}{s-1} + \frac{R+1}{s^{R+1}-1} \qquad (8)$$

The average distance from the president in an organization is minimal among all organizations of given s and Q. The reason for this is that R is minimal while $\frac{1}{s-1}$ is small and $\frac{R+1}{s^{R+1}-1}$ is negligible.

However, in the set of all balanced organizations, the regular organization has a maximal average distance between president and organization members. The reason is that the number of organization members at large distances is maximized. In fact, it is maximal, and not minimal, for a regular organization among all balanced organizations of radius R. Moreover, with a given bound on this span of control, average

distance from the president in a regular organization is maximal among all organizations of given s and given diameter (maximal distance between members) D.

In Section 10, we will consider information filtering and transmission in a regular organization. We may anticipate the result that average distance to the president being minimal, the amount of information lost in transmission to the president is also minimal, and so is the average loss of control by the president.

8.1 Average Wage and Unit Labor Cost. So far the input into an organization was considered in terms of persons only. As soon as money cost is considered, one must recognize the fact that personnel cost depends on personnel rank. Setting aside for the moment problems of motivation and qualification, economy of labor cost is economy of rank. An arrangement that minimizes rank for each person also minimizes labor cost. In the following only labor costs are considered. This is justified when the wage bill dominates all other costs or when other costs are a well-defined monotone increasing function of personnel cost such as when total cost is proportional to labor cost. For simplicity we shall describe personnel cost as wages. This must be taken to include all forms of compensation and the money equivalent of other perquisites and fringe benefits as well.

In any organization without slack there is a simple relationship between average wage and unit labor cost.

We assume that wages depend only on rank, $w = w_r$.

Let
$$W = \sum_{r=0}^{R} n_r w_r$$

be the total wage bill. Then unit labor cost

$$\bar{c} = \frac{W}{Q}.$$

Average wage is

$$\bar{w} = \frac{W}{N}.$$

In Section 5, it was shown that in organizations without slack

$$N = \frac{sQ-1}{s-1} \quad \text{or}$$

$$Q = \frac{1+(s-1)N}{s}$$

Therefore,

$$\bar{w} = \frac{W}{N} = \frac{W(s-1)}{Qs-1} = \frac{\bar{c}\cdot(s-1)}{s - \frac{1}{Q}} \tag{9}$$

$$\bar{c} = \frac{W}{Q} = \frac{W\cdot s}{1+(s-1)N} = \frac{\bar{w} \cdot s}{s-1 + \frac{1}{N}}. \tag{10}$$

Equation (9) and (10) state the relationship between average wage and unit labor cost. For large organizations $N >> 1$ or $Q >> 1$ one has approximately

$$\bar{c} \approx = \frac{\bar{w}}{1 - \frac{1}{s}} \tag{11}$$

Thus unit labor cost is approximately

$$\frac{1}{1 - \frac{1}{s}} \quad \text{times average wage.}$$

This relationship holds regardless of the ranks of operatives. The factor $\frac{1}{1 - \frac{1}{s}}$ was introduced in Chapter 4 as the organizational multiplier.

8.2 Minimizing Labor Cost

Consider an organization with given task Q. Assume that personnel can be hired in fractional units. We want to determine the organization with the minimal wage bill.

$$\underset{}{\text{Min}} \quad \sum_{r=0}^{R} w_r n_r$$

Recall from (7.6) that the presidential rank is at least equal to $\frac{\log Q}{\log s}$. We may write

$$R = \left\{\frac{\log Q}{\log s}\right\} = R(Q)$$

as a lower bound for R where $\{x\}$ denotes the smallest integer $\geqq x$.

The minimand is, therefore, at least equal to

$$\sum_{r=0}^{R(Q)} w_r n_r$$

The constraints imposed on organizational structure by the requirements of supervision are once more

$$N_{r+1} \geqq \frac{1}{s} (N_r - 1) \tag{4}$$

Moreover, N_0 is determined by the size of the task

$$N_0 = N = \frac{sQ-1}{s-1} \tag{12}$$

(Here the assumption of no slack, i.e., of fractional hiring is essential.) Rewrite the minimand in terms of N_r to obtain the following minimization problem

$$\underset{N_r}{\text{Min}} \quad \sum_{r=0}^{R(Q)} w_r (N_r - N_{r+1})$$

subject to

$$N_{r+1} \geqq \frac{1}{s} (N_r - 1) \qquad r - 1, \ldots, R-1 \tag{4}$$

N_0 given.

Rearranging terms in the minimand

$$\underset{N_r}{\text{Min}} \quad w_R n_R + \sum_{r=1}^{r-1} N_r (w_r - w_{r-1}) - w_0 N_0$$

Since wages increase with rank

$$w_r > w_{r-1} \qquad r = 1, \ldots, R$$

and N_0 is given, the minimand is an increasing function of every N_r $r = 1, \ldots, R$. Now inequality (4) may be applied repeatedly to yield lower bounds on each N_r. Thus

$$N_1 \geqq \frac{1}{s} N_0 - \frac{1}{s}$$

$$N_2 \geqq \frac{1}{s^2} N_0 - \frac{1}{s} - \frac{1}{s^2} \tag{4a}$$

$$\vdots$$

$$N_r \geqq \frac{1}{s^r} N_0 - \frac{1}{s} - \frac{1}{s^2} - \ldots - \frac{1}{s^r}$$

Clearly the minimand is minimized when the "=" sign applies in all inequalities (4a). Taking first differences of

$$N_r = \frac{1}{s^r} N_0 - \frac{1}{s} - \frac{1}{s^2} - \ldots - \frac{1}{s^r}.$$

one has

$$n_r = N_r - N_{r+1}$$

$$= N_0 \left(\frac{1}{s^r} - \frac{1}{s^{r+1}}\right) + \frac{1}{s^{r+1}}$$

$$= \frac{sQ-1}{s-1} \cdot \frac{s-1}{s^{r+1}} - \frac{1}{s^{r+1}} \qquad \text{using} \quad (12),$$

$$n_r = \frac{Q}{s^r} \tag{13}$$

This shows that the minimum is attained by a regular organization.

The numbers of positions n_r in rank r are intergers if and only if

$$Q = s^R.$$

We have shown:

Lemma 1:
Among all organizations with a given span of control s and a given task size Q, the regular organization has the smallest wage bill.

Corollary: The regular organization has the smallest unit labor cost and the smallest average wage among all organizations with the same task Q and span of control s.

To illustrate this point, consider all organizations using integer numbers of supervisors and operatives when $s = 3$ and $Q = 1, \ldots, 10$.

Table 7.1 Comparison of Unit Labor Costs

$Q = n_0$	1	2	3	4	5	6
n_1	0	1	1	2	3	2
n_2	0	0	0	1	1	1
$\bar{w}$	w_0	w_1+2w_0	w_1+3w_0	$4w_0+2w_1+w_2$	$5w_0+2w_1+w_2$	$6w_0+2w_1+w_2$
$\bar{c}$	w_0	$w_0+\frac{w_1}{2}$	$w_0+\frac{w_1}{3}$	$w_0+\frac{w_1}{2}+\frac{w_2}{4}$	$w_0+\frac{2}{5}w_1+w_2$	$w_0+\frac{1}{3}w_1+\frac{w_2}{6}$

	7	8	9	10
	3	3	3	3
	1	1	1	1
1	$7w_0+3w_1+w_2$	$8w_0+3w_1+w_2$	$9w_0+3w_1+w_2$	$10w_0+3w_1+w_2+w_3$
	$w_0+\frac{3}{7}w_1+\frac{w_2}{7}$	$w_0+\frac{3}{8}w_1+\frac{w_2}{8}$	$w_0+\frac{1}{3}w_1+\frac{w_2}{9}$	$w_0+\frac{3}{10}w_1+\frac{w_2}{10}+\frac{w_3}{10}$

Clearly unit labor cost is minimal for $Q = 1,3$ and 9, all representing regular organizations.

The proposition and corollary just proved may be generalized as follows.

Lemma 3: Among all organizations whose task is at least Q or whose size is at least N, the regular organization has the smallest unit cost and the smallest average wage.

Proof:

Let the constraint (4) be rewritten as

$$\frac{N_{r+1}}{N} \geq \frac{1}{s}\left(\frac{N_r}{N} - \frac{1}{N}\right) \tag{4b}$$

Repeated application yields

$$\frac{N_1}{N} \geqq \frac{1}{s} - \frac{1}{sN}$$

$$\frac{N_2}{N} \geqq \frac{1}{s^2} - \frac{1}{N}\left(\frac{1}{s} + \frac{1}{s^2}\right)$$

$$\vdots$$

$$\frac{N_r}{N} \geqq \frac{1}{s^R} - \frac{1}{N}\left(\frac{1}{s} + \ldots + \frac{1}{s^r}\right) \tag{14}$$

The right-hand side of the constraints (14) is an increasing function of N, tightening the constraints. Therefore, the minimum of the objective function

$$w = \sum_{r=0}^{R(Q)} w_r(N_r - N_{r+1})$$

is itself an increasing function of N. This proves the lemma for all $N \geqq N_0$, N_0 fixed.
To prove the lemma for $Q \geqq Q_0$ assume without restriction that all operatives have rank zero

$$N = N_1 + Q$$

$$N_1 \geqq \frac{1}{s}(N_1 + Q - 1)$$

from which

$$N_1 \geqq \frac{1}{s-1}(Q-1) \tag{15}$$

Next

$$N_2 \geqq \frac{1}{s}(N_1 - 1)$$

$$\geqq \frac{1}{s(s-1)} Q - \frac{1}{s-1} - \frac{1}{(s-1)s}$$

$$N_r \geqq \frac{1}{(s-1)s^{r-1}} Q - \frac{1}{s-1}\left(1 + \frac{1}{s} + \ldots + \frac{1}{s^{r-1}}\right)$$

Divide by Q

$$\frac{N_r}{Q} \geqq \frac{1}{(s-1)s^{r-1}} - \frac{1}{Q} \cdot \frac{1}{s-1}\left(1 + \frac{1}{s} + \ldots + \frac{1}{s^{r-1}}\right)$$

The right-hand side increases and the constraints become tighter with increasing Q. This proves the lemma.

9. Costs and Scale

To study the effect of scale on unit cost, it is necessary to specify a wage structure. In the following we consider only wage structures which are characterized by constant proportional increments resulting in

$$w_r = w_0 b^r \qquad b > 1 \qquad r = 0,1, \ldots, R \tag{1}$$

Such a wage scale is recommended already in ARTA SHASTRA [Kautiliya, (cf. Beckmann, 1978, cf. also Simon 1965].

Later on, we shall allow deviations from this at the top and bottom levels. Moreover, many conclusions remain valid when the organization's wage structure is dominated by a regular scale (1)

$$w_r \leqq w_0 b^r \tag{1a}$$

9.1 Regular Organization. The wage bill of a regular organization is easily calculated as

$$W = \sum_{r=0}^{R} w_r n_r = \sum_{r=0}^{R} w_o b^r s^{R-r} \quad \text{using (1) and} \quad (8.1)$$

$$= w_o s^R \sum_{r=0}^{R} \left(\frac{b}{s}\right)^r$$

$$W = w_o s^R \frac{1 - \left(\frac{b}{s}\right)^{R+1}}{1 - \frac{b}{s}} \tag{2}$$

using the well-known formula for the geometric series

$$1 + a + a^2 + \ldots + a^n = \frac{1 - a^{n+1}}{1-a} \qquad a \neq 1 \tag{3}$$

An alternative expression for (2) is

$$w = w_0 \frac{s^{R+1} - b^{R+1}}{s-b} \tag{2a}$$

From (2) one may calculate the average wage $\bar{w}$ and the unit labor cost $\bar{c}$. Observe first that

$$N = 1 + s + \ldots + s^R = \frac{s^{R+1}-1}{s-1} \tag{4}$$

Now

$$\bar{w} = \frac{W}{N} = w_0 \frac{s-1}{s-b} \cdot \frac{1 - (\frac{b}{s})^{R+1}}{1 - (\frac{1}{s})^{R+1}} \tag{5}$$

Since $b > 1$ this is an increasing function of R (as asserted by Lemma 3 of Section 8). It approaches the limit

$$\bar{w} \to w_o \frac{s-1}{s-b} \tag{6}$$

Of greater importance for economic theory is the relationship between unit cost and scale.

$$\bar{c} = \frac{W}{Q} = \frac{W}{s^R}$$

$$\bar{c} = \frac{w_0}{1 - \frac{b}{s}} \cdot [1 - (\frac{b}{s})^{R+1}] \tag{7}$$

Once more this is an increasing and bounded function of top rank R and hence of the scale of the organization, whether measured in terms of R, Q or N. Its limit is

$$\bar{c} \to \frac{w_0}{1 - \frac{b}{s}} \tag{8}$$

It is proportional to the basic wage w_0 and is an increasing function of the growth factor b for wages and a decreasing function of the span of control s, as one would expect.

9.2 Extended Regular Organization. How are these simple relation-ships between cost and scale affected when wages and span of control are subject to modification at the top and bottom levels? Assume

$$\begin{aligned} n_R &= 1 \\ n_{R-1} &= s_R \\ &\vdots \\ n_r &= s_R\, s^{R-r-1} \\ n_1 &= s_R s^{R-2} \\ n_0 &= s_R s^{R-2} s_1 \end{aligned} \tag{9}$$

Typically (but not necessarily)

$$s_R < s < s_1$$

Moreover, assume

$$w_r = w_1 b^{r-1} \qquad r = 1, \ldots, R-1 \tag{10}$$

allowing w_0 and w_R to be exceptional. A straightforward calculation yields

$$W = \sum_{r=0}^{R} w_r n_r$$

$$= w_0 s_R s^{R-2} s_1 + \sum_{r=1}^{R-1} w_1 b^{r-1} s_R s^{R-r-1} + w_R$$

$$= w_0 s_R s^{R-2} s_1 + w_1 s_R \frac{s^{R-1} - b^{R-1}}{s-b} + w_R$$

$$= [w_0 + \frac{w_1}{s_1(s-b)}] s_R s^{R-2} s_1 + w_R + w_R - w_1 \frac{s_R}{s-b} b^{R-1} \quad (12)$$

Unit cost is now

$$\bar{c} = \frac{W}{Q} = \frac{W}{s_R s^{R-2} s_1}$$

$$\bar{c} = w_0 + \frac{w_1}{s_1(s-b)} + \frac{w_R - w_1 \frac{s_R}{s-b} \cdot b^{R-1}}{s_R s^{R-2} s_1} \quad (13)$$

Notice first that unit cost is constant throughout provided presidential compensation is set at the level

$$w_R = w_1 \frac{s_R}{s-b} \cdot b^{R-1}$$

The wage $w_1 b^{R-1}$ would be in line with wages of other managers. When $s_R > s - 1$ the formula (14) for presidential compensation would yield larger values.

In the following we suppose that presidential compensation is still a function of rank R of the simple form

$$w_R = b_0 b^{R-1} w_1 \quad (15)$$

when typically $b_0 > 1$. Then

$$\bar{c} = w_0 + \frac{w_1}{s_1(s-b)} + \frac{(b_0 - \frac{s_R}{s-b}) b^{R-1}}{Q} \quad (16)$$

Once more this is a bounded function, increasing when

$b_0 < \frac{s_R}{s-b}$, and approaching rapidly its limiting value

$$\bar{c} + w_0 + \frac{w_1}{s_1(s-b)} \tag{17}$$

Thus in the limit unit labor cost consists of direct labor cost

w_0 and the prorated cost of administration $\frac{w_1}{s_1(s-b)}$.

9.3 Quasi-Regular Organization. A quasi-regular organization is defined as an organization with fixed span of control, in which a fixed proportion of positions at every level (except the bottom level) is nonsupervisory. The simplest example is that of Figure 8.1 where only the three top ranks are shown.

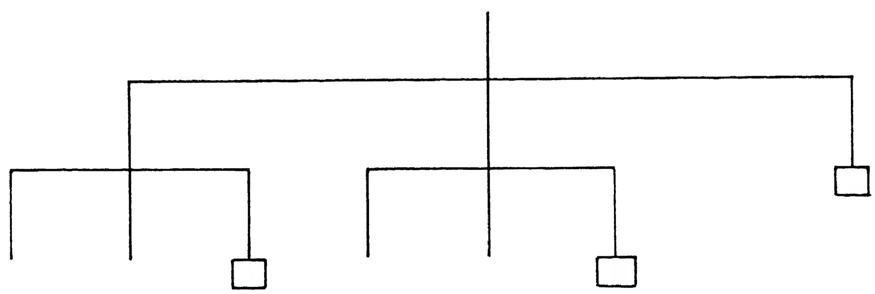


Figure 8
Nonsupervisory Positions in a Quasi-Regular Organization

In this case

$$n_R = 1$$

$$n_{R-1} = s$$

$$n_{R-2} = s(s-1)$$

$$\vdots$$

$$n_r = s(s-1)^{R-r-1} \qquad r = 0, \ldots, R-1 \tag{18}$$

Assume once more the regular wage scale (1). The wage bill is now

$$W = w_0 b^R + \sum_{r=0}^{R-1} w_0 b^r s(s-1)^{R-r-1}$$

$$W = w_0 b^R + w_0 s(s-1)^{R-1} \frac{1 - (\frac{b}{s-1})^R}{1 - \frac{b}{s-1}}$$

Consider once more unit cost as a function of scale R

$$\bar{c} = \frac{W}{Q} = \frac{W}{s(s-1)^{R-1}}$$

$$\bar{c} = \frac{w_0}{1 - \frac{b}{s-1}}$$

$$- w_0 (\frac{b}{s-1})^R \cdot \frac{1+b}{s-1-b} \cdot \frac{s-1}{s} \qquad (19)$$

This is a decreasing bounded function of scale approaching rapidly the limiting value

$$\bar{c} \to \frac{w_0}{1 - \frac{b}{s-1}} \qquad (20)$$

<u>Problem</u>. Consider a quasi-regular organization in which the following pattern for nonsupervisory positions is repeated throughout

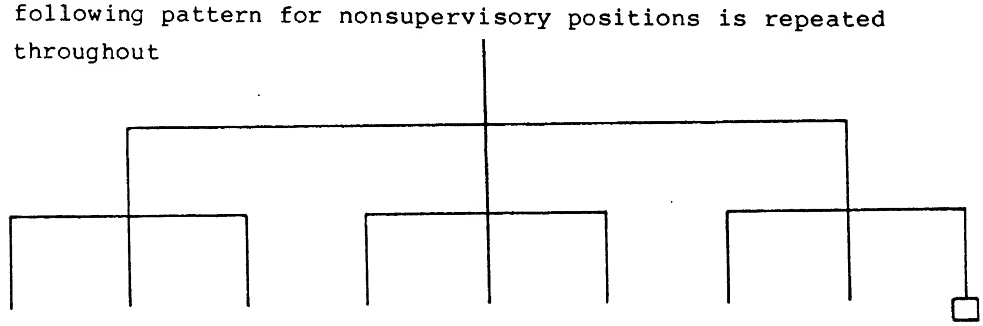


Figure 9 A Quasi-Regular Organization

Show that $n_r = \begin{cases} (s-1)^k s^k & \text{for } R-r = 2k \\ (s-1)^k s^{k+1} & \text{for } R-r = 2k+1 \end{cases}$

and derive expressions for the wage bill and average cost. Notice that the 'effective span of control' in this organization is $\sqrt{(s-1)s}$.

9.4 Cost and Scale for Nonregular Organizations.

In a quasi-regular organization, the relationship between unit labor cost and scale applies approximately as in a regular organization with an "effective span of control $\tilde{s} < s$ substituted for s.

Consider, however, the simple type of organization, described by Figure 7, an s ladder:

$$s_R = 1$$
$$s_r = s \qquad r\ 0,1,\ \ldots,\ R-1$$

There is only one supervisory position at every level. The number of operatives is

$$Q = s + (R-1)(s-1) = 1 + R \cdot (s-1) \tag{21}$$

the total number of positions is

$$N = 1 + sR$$

and the wage bill is

$$W = w_R + s(w_{R-1} + w_{R-2} + \ldots + w_1 + w_0) \tag{22}$$

Therefore, average wages are

$$\bar{w} = \frac{w_R + s(w_{R-1} + \ldots + w_0)}{1 + sR} \tag{23}$$

and unit labor costs are

$$\bar{c} = \frac{w_R + s(w_{R-1} + \ldots + w_0)}{1 + (s-1)R} \tag{24}$$

Consider the unweighted mean wage

$$\bar{w}_R = \frac{w_0 + w_1 + \ldots + w_R}{R + 1} \tag{25}$$

In terms of this the average wage (23) may be written

$$\bar{w} = \frac{w_R + sR\bar{w}_{R-1}}{1 + sR}$$

$$> \bar{w}_{R-1}$$

provided

$$w_R > w_{R-1} > \ldots > w_1 > w_0. \tag{26}$$

If the wage sequence w_r is unbounded, it follows that average wage in an s-ladder type of organization increases without bound.

Similarly consider unit labor cost

$$\bar{c} = \frac{w_R + sR\bar{w}_{R-1}}{1 + (s-1)R}$$

$$> \frac{s}{s-1}\, \bar{w}_{R-1}$$

provided (6) applies. Once more unit labor cost increases without bound when wages w_r increase without bound.

Specifically when

$$w_r = b^r \tag{27}$$

the unweighted mean wage is

$$\bar{w}_{R-1} = \frac{1 + b + \ldots b^{R-1}}{R} = \frac{b^R - 1}{R(b-1)} \tag{28}$$

and since $b > 1$ this lower bound increases with R without limits.

9.5 Regular Organization with Staff

Suppose that every line supervisor has a staff headed by one staff person whose rank is one below this supervisor's rank. Let the span of control for staff be $u \leq s$. Write

x_r supervisors of rank r

z_r staff of rank r

$n_r = z_r + x_r$ personnel of rank r.

Then

$$z_r = uz_{r+1} + x_{r+1} \tag{29}$$

Let s be the number of subordinates excluding staff per supervisor. Then

$$x_R = 1 \qquad z_R = 0$$

$$x_{R-1} = s \qquad z_{R-1} = 1$$

$$x_{R-2} = s^2 \qquad z_{R-2} = u + s$$

$$x_r = s^{R-r} \qquad z_r = u^{R-r-1} + u^{R-r-2}s + \ldots + us^{R-r-2} + s^{R-r-1}$$

$$z_r = \begin{cases} \dfrac{s^{R-r} - u^{R-r}}{s-u} & \text{if } u \neq s \\ (R-r)s^{R-r-1} & \text{if } u = s. \end{cases} \tag{30}$$

A regular organization with staff has the following characteristics:

Highest rank or "height" $\quad R$

Operatives or output $\quad Q = s^R \qquad (31)$

Total number of supervisors

$$X = 1 + s + s^2 + \dots s^R = \frac{s^R - 1}{s-1} \approx \frac{1}{s-1} Q \qquad (32)$$

Total staff

$$Z = \begin{cases} \dfrac{1}{s-u} \left[\dfrac{s^{R+1}-1}{s-1} - \dfrac{u^{R+1}-1}{u=1}\right] & u \neq s \\[2ex] \dfrac{Rs^R}{s-1} - \dfrac{s^R - 1}{(s-1)^2} & u = s \end{cases} \qquad (33)$$

Total size

$$N = Q + X + Z$$

Consider first $u < s$

$$N = \frac{sQ-1}{s-1} + \frac{1}{s-u} \left[\frac{sQ-1}{s-1} - \frac{u^{R+1}-1}{u-1}\right]$$

$$= \left(\frac{sQ-1}{s-1}\right) \cdot \left(1 + \frac{1}{s-u}\right) - \frac{u^{R+1}-1}{(s-1)(s-u)}$$

For large Q this is approximately proportional to Q

$$\frac{N}{Q} \doteq \frac{s}{s-1} \left(1 + \frac{1}{s-u}\right) = \frac{1 + \dfrac{1}{s-u}}{1 - \dfrac{1}{s}} \qquad (34)$$

Now let $u = s$,

$$N = \frac{sQ-1}{s-1} + \frac{Rs^R}{s-1} - \frac{s^R-1}{(s-1)^2}$$

$$= \frac{(R+s)Q-1}{s-1} - \frac{Q-1}{(s-1)^2}$$

For large Q approximately

$$\frac{N}{Q} \doteq \frac{R+s}{s-1} - \frac{1}{(s-1)^2} \tag{35}$$

When the staff span of control u equals the line span of control s, then the ratio of total positions to operative positions is a linear increasing function of R.
Recall that R in turn is a logarithmic function of Q

$$R \doteq \frac{\log Q}{\log s} \tag{36}$$

When $u < s$, consider next the wage bill

$$W = w_0 \cdot \frac{1 + \frac{1}{s-u}}{1 = \frac{b}{s}} Q$$

with approximate equality when Q is large. Hence unit labor cost approximates the constant value

$$\bar{c} \to \frac{1 + \frac{1}{s-u}}{1 - \frac{b}{s}} \tag{37}$$

When $u = s$

$$n_r = s^{R-r} + (R-r)s^{R-r-1} \qquad r = 0, \ldots, R-1$$

For $w_r = w_0 b^r$ the wage bill equals

$$W = w_0 \sum_{r=0}^{r-1} b^r [s^{R-r} + (R-r)s^{R-r-1}] + w_0 b^R$$

$$= w_0 \left(\frac{s^{R+1} - b^{R+1}}{s-b} + \frac{Rs^R}{s-b} - \frac{bs^R}{(s-b)^2} + \frac{1}{(s-b)^2}\right)$$

Hence, using $Q = s^R$,

$$\bar{c} = \frac{W}{Q} = \frac{w_0}{1 - \frac{b}{s}} + \frac{w_0 (\frac{b}{s})^{R+1}}{1 - \frac{b}{s}}$$

$$+ \frac{w_0 R}{s-b} - \frac{w_0 b}{(s-b)^2} + \frac{w_0}{(s-b)^2} \quad \frac{1}{Q}$$

For large Q this is approximately

$$\bar{c} = w_0 \left(\frac{s}{s-b} - \frac{b}{(s-b)^2} + \frac{R}{s-b}\right) \tag{38}$$

Hence, unit labor cost increases linearly with R.

The conclusion to be drawn from these examples is as follows. Unless an organization is regular or quasi-regular, unit labor cost will increase with scale without limit. Such organizations exhibit decreasing returns (or increasing unit cost) to scale.

10. Loss of Communication and Control

10.1 Communication

The higher is his/her rank, the closer a person is to the "center of things" in an organization, particularly for information originating with the president. Also the more removed is a person from information input by nonsupervisors (operatives).

If major decisions issue from the president's office then proximity to the center, measured by $R-r$ in counting down determines how soon and sometimes how much you know.

Filtering or dilution of information at some loss rate δ turns out to be necessary to prevent the system from being flooded with information.

Consider an organization member i. From various other members j in a source set S_i, information μ_j (in bits, say) is received per unit of time, totalling $\sum_{j \varepsilon S_i} \mu_j$.

This is filtered and processed yielding an amount of information

$$(1 - \delta_i) \sum_{j \varepsilon S_i} \mu_j$$

In general, further information m_i is added. Information outflow is then

$$\mu_i = m_i + (1-\delta_i) \sum_{j \varepsilon S_i} \mu_j \qquad (1)$$

and this is passed to all members k in a recipience set T_i, $k \in T_i$ continuing the process.

For all members i this flow must be kept below capacity, say unity $\mu_i \leqq 1$. When official channels are used this can be made more specific. Let each supervisor have no more then s subordinates directly reporting to him/her and let each member filter the information received at a loss rate of at least δ and generate additional information at a rate not exceeding m. Then the rate μ_r of emission of information by a member of rank r in terms of the rate μ_{r-1} of information received from the s subordinates is restricted by

$$\mu_r \leqq m + (1-\delta)s\mu_{r-1}. \qquad (2)$$

Through successive substitution

$$\mu_r \leqq m \cdot [1 + (1-\delta)s + (1-\delta)^2 s^2 + \ldots + (1-\delta)^r s^r]$$

$$= m \; \frac{1 - (1-\delta)^{r+1} s^{r+1}}{1 - (1-\delta)s}$$

or

$$\mu_r < \frac{m}{1 - (1-\delta)s} \qquad (3)$$

A sufficient condition for this upward information flow to lie within the capacity 1 for information emission of any member is

$$\frac{m}{1 - (1-\delta)s} < 1$$

This may be expressed in three different ways.

$$m \leqq 1 - (1-\delta)s \qquad (4)$$

or

$$\delta \geqq 1 - \frac{1-m}{s} \qquad (5)$$

or

$$s \leqq \frac{1-m}{1-\delta} \qquad (6)$$

Constraints (4), (5), and (6) are equivalent. Condition (4) limits the amount of information that can be generated when δ and s are given; condition (5) imposes a lower bound on the elimination rate δ for information, when m and s are given. Finally, condition (6) limits the span of control, when information generation m and elimination rate δ are given.

10.2 Loss of Control

As suggested by Williamson [1975] the effectiveness of labor is reduced by a factor of ρ^R, $0 < \rho < 1$ when top management is separated from labor by a chain of R supervisors. Thus output Q and labor force L are related by

$$Q = \rho^R L \qquad (7)$$

In a regular organization with span of control s, size N and amount of labor L were shown to be related by

$$N = \frac{sL-1}{s-1}$$

$$= \frac{s\rho^{-R}Q-1}{s-1} \qquad \text{using (7),}$$

or dropping the 1

$$N \doteq \frac{\rho^{-R}}{1 - \frac{1}{s}} \cdot Q \qquad (8)$$

To determine R observe that a constant span of control s implies

$$s^R = L = \rho^{-R}\, Q \qquad \text{or} \tag{9}$$

$$Q = (\rho s)^R$$

$$R = \frac{\ln Q}{\ln (\rho s)} \tag{10}$$

Substitute (10) in (9)

$$N = \frac{Q}{1 - \frac{1}{s}} \; e^{-R \ln \rho}$$

$$= \frac{1}{1 - \frac{1}{s}}\, Qe^{-\ln Q \cdot \frac{\ln \rho}{\ln \rho s}}$$

$$= \frac{1}{1 - \frac{1}{s}}\; Q^{1 - \frac{\ln \rho}{\ln \rho s}}$$

$$N = \frac{1}{1 - \frac{1}{s}}\; Q^{\eta} \tag{11}$$

where

$$\frac{1}{\eta} = 1 + \left| \frac{\ln \rho}{\ln \rho s} \right| \tag{12}$$

Observe that $\eta > 1$
Solving (12) for Q

$$Q = a\, N^{\eta} \quad \text{with} \quad \eta < 1 \tag{13}$$

This is a Cobb-Douglas production function for output considered as a function of total personnel N as input. Notice that an exponent $\eta < 1$ would be consistent with the presence of another

fixed factor, a hidden "management" component. In other words the Williamson formula of control loss is formally identical with and cannot be distinguished from a Cobb-Douglas production function.

IV. Organizational Production Functions

11. Case Working Organizations

We are thus led to reconsider the relationship between labor and management inputs and organizational output in terms of the familiar concept of a production function (Beckmann, 1977, 1982, Sato, 1981).

11.1 Introduction. In understanding the economics of organizations the notion of a production function will prove useful. Its inputs are operative labor and managerial labor at various levels of supervision and the output a measure of the rate at which the task of the organization is discharged [Beckmann, 1977]. While this may seem a natural way of thinking to those indoctrinated by microeconomic theory, some demonstration appears to be in order, how such a relationship may arise in which labor and management appear as factors of production. In the words of one reviewer of an NSF proposal "it is possible that the organizational production function is a useful tool for some purposes, but I really would like to see some more microeconomics before rushing to write down a Cobb-Douglas production function, one of whose arguments is supervision."

The relationship between inputs and outputs in a hierarchically structured organization has been studied as a cost function in Kochen-Deutsch [1974] with communication treated as a decision variable. The economics of hierarchical structures is attracting increasing attention, Williamson [1975], Mirrlees [1976]. Beckmann [1978], Calvo [1979].

The purpose of this section is a much narrower one: to develop a production function for organizations whose output may be measured in terms of "cases handled". The analysis uses elementary queuing theory. In view of the restrictiveness of this queuing model, the results should be considered as illustrative rather than definitive.

Examples of organizational tasks that can be described as cases are:

auditing of income tax returns,

adjusting insurance claims,

determining and monitoring the eligibility of welfare applicants.

There are, of course, many other instances where paper work can be divided into simple units to be labelled cases.

11.2 Model. To begin with we consider a simple organization consisting of $X_1 = 1$ supervisor and X_o caseworkers. Each case worker has a well-defined area of competence, determined by either geographical or technical criteria that are easy to verify. Cases arrive at random, and are assigned without delay to the appropriate caseworker by the supervisor's office. The areas of competence are delimited in such a way that each caseworker receives on the average the same case load per period. If cases arrive at random at a rate λ, then those assigned to a particular caseworker arrive at random at a rate λ.

Define the completion rate at which a caseworker handles cases to be μ_o and assume this to be the same for each caseworker. Alternatively we may say that the average time needed by a caseworker to complete a case is $\frac{1}{\mu_o}$.

Introduce now the rather special assumption that the completion times are exponentially distributed with mean $\frac{1}{\mu_o}$. Then it is well-known that cases will spend an average length of time

$$\frac{1}{\mu_o - \frac{\lambda}{x_o}} \qquad (1)$$

with the respective caseworker [Wagner, 1969]. When a caseworker has finished a case, it is then passed back to the supervisor for possible inspection, verification and signing. The supervisors' completion rate is defined to be μ_1, that is, the average time required for inspection verification and signing is $\frac{1}{\mu_1}$. Once more assume that this time is exponentially distributed with mean $\frac{1}{\mu_1}$. A case will then spend an average time

$$\frac{1}{\mu_1 - \frac{\lambda}{x_1}} \qquad (2)$$

while waiting for and being disposed of by the supervisor.

Although we consider $x_1 = 1$, as the normal situation, the possibility of the supervisor being available only part-time ($x_1 < 1$) is included here as a possibility. For the waiting time formula to be valid, the times of availability must be randomly chosen.

Sometimes the case will be returned to the caseworker and possibly there will be some interaction during the actual casework between worker and supervisor. All this is implicit in the (exponential) probability time distributions for casework and, therefore, need not be considered here explicitly.

The expected value of the total time that a case spends in the system on the average is now the sum of expressions (1) and (2)--a special property of exponential service times distributions. We will now introduce the important postulate that this average time must not exceed an acceptable level τ in a well-functioning organization.

In welfare agencies operating at the county level, the State of California imposes in fact a time limit on eligibility

procedures as a condition for its financial support. This time limit is meant to apply in every case, but in the face of randomness, this is clearly not a realistic requirement.

Let limit τ on average time in the system be effective

$$\frac{1}{\mu_o - \frac{\lambda}{x_o}} + \frac{1}{\mu_1 - \frac{\lambda}{x_1}} = \tau. \qquad (3)$$

Consider the system in steady-state equilibrium. The rate of case inflow λ is then also the rate of case outflow and this may be considered a measure of the output of the organization. The inputs are $x_o = 1$ supervisor and x_1 caseworkers.

Equation (3) is an implicit definition of a production function

$$\lambda = \lambda(x_o, x_1) \qquad (4)$$

in which output λ depends on two factors of production, labor x_o and supervision x_1.

11.3 Solution. We solve (3) in closed form and establish that it has the usual properties of a microeconomic production function.

Straightforward arithmetic yields the following quadratic equation in λ

$$\lambda^2 - [ax_o + bx_1]\lambda + [ab - \frac{1}{\lambda^2}]x_o x_1 = 0$$

where

$$a = \mu_o - \frac{1}{\tau} > 0$$

$$b = \mu_1 - \frac{1}{\tau} > 0.$$

That a and b are positive is seen as follows. Notice that τ is feasible only if it exceeds the requirements $\frac{1}{\mu_o}$ and $\frac{1}{\mu_1}$ at each stage for processing

$$\tau > \frac{1}{\mu_i} \qquad i = 0,1.$$

This equation is solved by

$$\lambda = \frac{ax_o - bx_1}{2} - \frac{1}{2} \sqrt{(ax_o - bx_1) + \frac{4x_o x_1}{\tau^2}}. \qquad (5)$$

Notice that

$$\lambda(0,x_1) = 0$$

implies the minus sign for the square root term.

As a first property of the production function (5), we notice that output increases with τ, when x_o, x_1 are held constant. This follows by differentiation of (3)

$$\frac{\partial\tau}{\partial\lambda} = \frac{1}{(\mu_o - \frac{\lambda}{x_o})^2 x_o} + \frac{1}{(\mu_1 - \frac{\lambda}{x_1})^2 x_1} > 0.$$

τ may be considered a quality attribute of output; as τ increases quality decreases. Thus increasing quality decreases quantity of output for given inputs x_o, x_1.

That output λ is an increasing function of either input x_i is best seen from (3). The left-hand expression increases with λ and decreases with x_i so that by the implicit function theorem

$$\frac{\partial\lambda}{\partial x_i} = - \frac{\frac{\partial\tau}{\partial x_i}}{\frac{\partial\tau}{\partial\lambda}}$$

$$\text{sign } \frac{\partial\lambda}{\partial x_i} = - \frac{-}{+} = + \qquad (6)$$

Thus the marginal product of either factor is positive. From (3) it may be shown that $\lambda(x_o, x_1)$ is linear homogeneous, but this is also apparent from (5). In fact, writing

$$u = \frac{x_i}{x_o} \qquad v = \frac{\lambda}{x_o} \tag{7}$$

one has the homogenized production function in one variable

$$v(u) = \frac{a}{2} + \frac{b}{2}u - \frac{1}{2}\sqrt{(a-bu)^2 + \frac{4u}{\tau^2}} \tag{8}$$

For λ to be concave, i.e., for the law of diminishing returns to apply, it is sufficient that $v(u)$ be concave. Now a straightforward calculation shows that

$$\text{sign } v''(u) = \text{sign } \left(\frac{1}{\mu_o} + \frac{1}{\mu_1} - \tau\right)$$
$$< 0 \quad \text{if} \quad \tau \text{ is feasible.} \tag{9}$$

Recall that the $\frac{1}{\mu_i}$ are the expected processing times in stage i for each case.

As a final property of the production function λ, we note that for $x_1 = 1$ (or in fact $x_1 = $ constant) the output is bounded. For letting $x_o \to \infty$ in (3) one obtains

$$\frac{1}{\mu_o} + \frac{1}{\mu_1 - \lambda} \leqq \tau$$

which is solved by

$$\lambda(x_o, 1) < \mu_1 - \frac{1}{\tau - \frac{1}{\mu_o}}$$

for any finite x_o. Similarly for fixed x_o, the production function is bounded with respect to the variable x_1.

As an illustration consider the case with the simplest numbers

$\mu_o = \mu_1 = 3 \quad \tau = 1$ satisfying (9) and implying

$a = b = 2,$

$$\lambda = x_o + x_1 - \sqrt{x_o^2 - x_o x_1 + x_1^2} \qquad (11)$$

By a linear transformation of the variables x_o, x_1, λ the function λ can always be standardized as (11). The homogenized version of (11)

$$v = 1 + u - \sqrt{1 - u + u^2} \qquad (12)$$

is illustrated in Figure 10.

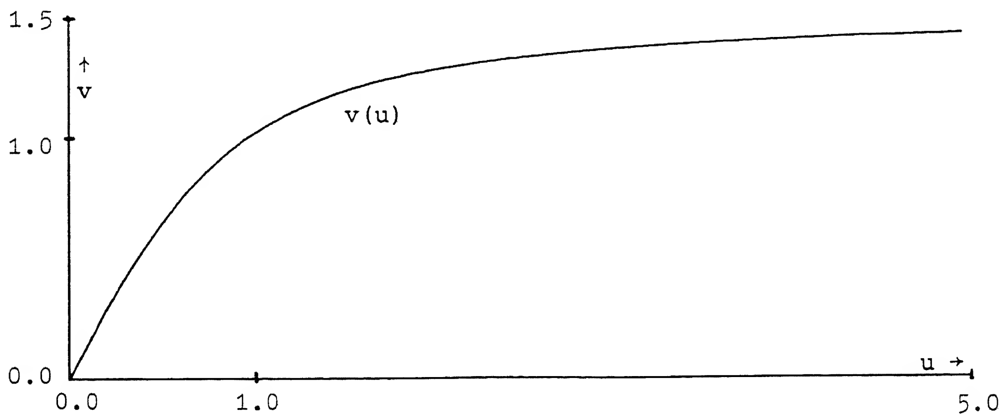


Figure 10. Production Function

Boundedness of the production function implies that it is not well approximated by such standard types as a Cobb-Douglas function for large values of x_o, given $x_1 = 1$. Rather the production functions (11), (12) must stand on their own.

11.4 Discussion. It should be pointed out that in the present model the output rate μ_o of a caseworker does not depend on the amount of supervision $\frac{x_1}{x_o}$ that might be calculated to fall on one

caseworker. It is not that number which is relevant, but the fact that each case goes to the supervisor for inspection and that the necessary inspection time averaging $\frac{1}{\mu_1}$ is always applied. Thus the fact of supervision implies the necessary amount of supervision. The output rate $\frac{1}{\mu_o}$ of a caseworker is often set by general rules in an organization. Thus in welfare agencies, case load per worker is prescribed. In one California agency actual case loads for eligibility work were as follows

633/33 = 20.1 for new applications

6,250/64 = 97.6 for continuing cases.

Since continuing cases are to be rechecked every month, a case load of, e.g., 6,250 cases should be interpreted as a flow of 6,250 cases per month. In the budget proposal it was stated that these case loads were within the limits set by the state (these limits may be inferred as 20 and 100, respectively).

One important function of the supervisor's checking a case is to detect errors. The admissible "error rate" is another quality attribute of output. In welfare agencies this is also another variable on which the state imposes control as a condition of financial support. In the agency mentioned before, the actual "payment error rate" of 8.55% exceeded the state's limit of 4% during the report year.

Considering the error rate as a function of the amount of checking $\frac{1}{\mu_1}$ that is applied to a case on the average introduces another quality variable into the production function.

It can be argued that error detection is subject to the laws of search

Probability (undetected error) $\rho = e^{-\alpha/\mu_1}$

where α measures the effectiveness of search at $\frac{1}{\mu_1}$ is the average time spent on search so that

$$\mu_1 = \frac{\alpha}{|\log \rho|}$$

and

$$\frac{1}{\mu_o - \frac{\lambda}{x_o}} + \frac{1}{\frac{\lambda}{|\log \rho|} - \frac{\lambda}{x_1}} = \tau .$$

This production function which has the same functional form in terms of x_o and x_1 as before now contains the error rate ρ as a parameter. This point will not be pursued further here.

Instead of the average time spent by the case in the system one might control the probability ε that a case spends more than a prescribed time T in the system. Now the passage time at each stage is exponentially distributed, and the probability that their sum exceeds T equals

$$\frac{(\mu_1 - \frac{\lambda}{x_1})\, e^{-(\mu_1 - \frac{\lambda}{x_1})T} + (\mu_o - \frac{\lambda}{x_o})\, e^{-(\mu_o - \frac{\lambda}{x_o})T}}{\mu_1 - \mu_o - \frac{\lambda}{x_1} + \frac{\lambda}{x_o}} = 1 - \varepsilon . \tag{14}$$

This equation involves more complicated functional forms and two parameters T, ε instead of one, τ. The production function defined by (14) can be shown to have the same qualitative properties as (5): It is monotone increasing, concave, linear homogeneous and bounded. When

$$\mu_1 - \frac{\lambda}{x_1} >> \mu_o - \frac{\lambda}{x_o}$$

i.e., when delays at the supervisory level are small relative to those at the caseworker level, then one has approximately

$$\frac{\mu_1 - \frac{\lambda}{x_1}}{\mu_1 - \mu_o - \frac{\lambda}{x_1} + \frac{\lambda}{x_o}} \; e^{-(\mu_o - \frac{\lambda}{x_o})T} = \varepsilon. \qquad (15)$$

A Taylor expansion of the exponential function yields then once more a quadratic function in λ as a first approximation.

11.5 Generalization. In this analysis, we show how the assumption of exponential service time distributions can be dropped. As noted by Cox [1967] the convergence of several streams of customers, even after being processed with non-exponential service times, generates an approximate Poisson stream, if the original arrivals at the different caseworkers were independent and Poisson.

Let σ_i^2 $i = 0,1$ be the variances of the service time distributions for caseworkers 0 and supervisors 1 respectively. The Pollacek-Kintchin formula [Wagner 1965] states that average delay (including service time) at station i equals

$$W_i = \frac{1}{\mu_i} + \frac{(\frac{\lambda}{x_i\mu_i})^2 + (\frac{\lambda}{x_i})^2 \sigma_i^2}{2\frac{\lambda}{x_i}(1 - \frac{\lambda}{x_i\mu_i})}. \qquad (16)$$

Adding and simplifying we obtain the total average delay τ

$$\frac{\frac{\lambda}{x_o\mu_o^2} + \frac{\lambda\sigma_o^2}{x_o}}{2(1 - \frac{\lambda}{x_o\mu_o})} + \frac{\frac{\lambda}{x_1\mu_1^2} + \frac{\lambda\sigma_1^2}{x_1}}{2(1 - \frac{\lambda}{x_1\mu_1})} + \frac{1}{\mu_o} + \frac{1}{\mu_1} = \tau. \qquad (17)$$

Once more equation (17) defines a production function

$$\lambda = \lambda(x_o, x_1, \tau).$$

It may be found in closed form by solving the quadratic equation obtained from (2) by straightforward arithmetic

$$- \lambda^2 \left[\frac{1}{x_o x_1 \mu_o^2 \mu_1} + \frac{1}{x_o x_1 \mu_1^2 \mu_o} + \frac{\sigma_o^2}{x_o x_1 \mu_1} + \frac{\sigma_1^2}{x_o x_1 \mu_o} + \frac{2(\tau - \frac{1}{\mu_o} - \frac{1}{\mu_1})}{x_o x_1 \mu_o \mu_1} \right] + \lambda \left[\frac{1}{x_o \mu_o^2} + \frac{\sigma_o^2}{x_o} + \frac{1}{x_1 \mu_1^2} + \frac{\sigma_1^2}{x_1} + 2(\tau - \frac{1}{\mu_o} - \frac{1}{\mu_1})(\frac{1}{x_o \mu_o} + \frac{1}{x_1 \mu_1}) \right] = 2(\tau - \frac{1}{\mu_o} - \frac{1}{\mu_1})$$

Multiplying both sides by $- x_o x_1$ yields

$$a\lambda^2 - (bx_o + cx_1)\lambda = -dx_o x_1 \tag{18}$$

with

$$a = \frac{1}{\mu_o^2 \mu_1} + \frac{\sigma_o^2}{\mu_1} + \frac{1}{\mu_o \mu_1^2} + \frac{\sigma_1^2}{\mu_o} + 2\, \frac{\tau - \frac{1}{\mu_o} - \frac{1}{\mu_1}}{\mu_o \mu_1} > 0$$

$$b = \frac{1}{\mu_1^2} + \sigma_1^2 + 2(\tau - \frac{1}{\mu_o} - \frac{1}{\mu_1})\frac{1}{\mu_1} > 0$$

$$c = \frac{1}{\mu_o^2} + \sigma_o^2 + 2(\tau - \frac{1}{\mu_o} - \frac{1}{\mu_1})\frac{1}{\mu_o} > 0 \tag{19}$$

$$d = 2(\tau - \frac{1}{\mu_o} - \frac{1}{\mu_1}) > 0$$

Equation (18) is solved by

$$\lambda = \frac{bx_o + cx_1}{2a} \pm \frac{1}{2a} \sqrt{(bx_o+cx_1)^2 - 4adx_ox_1} \qquad (20)$$

By nature of a production function

$$\lambda(0,x_1) = \lambda(x_o,0) = 0$$

This implies the minus sign for the square root. In terms of standardized variables

$$q_o = \frac{b}{a}\, x_o \qquad\qquad q_1 = \frac{c}{a}\, x_1 \qquad (21)$$

equation (20) assumes the simple form

$$\lambda = \frac{q_o + q_1}{2} - \frac{1}{2} \sqrt{(q_o+q_1)^2 - 4\,\delta q_o q_1} \qquad (22)$$

This is a universal production function for case work in standardized form. It depends on only one parameter δ whose effects will now be studied. First observe that

$$\delta = \frac{ad}{bc} < 1 \qquad (23)$$

is necessary and sufficient for (22) to be real valued. It may be shown by a straightforward but tedious calculation that condition (19) is necessary and sufficient for the radicand in (22) to be positive for all q_o, q_1. Notice that $\lambda(q_o,q_1) \geqq 0$ and $=$ only when $q_o = 0$ or $q_1 = 0$. Moreover,

$$2\,\frac{\partial\lambda}{\partial q_i} = 1 - \frac{(q_o+q_1) - 2\,\delta q_j}{\sqrt{(q_o+q_1)^2 - 4\,\delta q_o q_1}} > 0 \quad \text{for} \quad \delta < 1 \qquad (24)$$

$$2\,\frac{\partial^2\lambda}{\partial q_i^2} = -\,(q_o+q_1)^2 - 4\delta q_o q_1 + \frac{(q_o+q_1 - 2\delta q_j)^2}{(q_o+q_1)^2 - 4\delta q_o q_1}$$

$$= \frac{4\delta(\delta-1)q_j^2}{(q_o+q_1)^2 - 4\delta q_o q_1} < 0 \quad \text{for} \quad \delta < 1 \tag{25}$$

so that the law of diminishing returns applies. Notice, however, that the Hessian is singular since λ is linear homogeneous in x_o, x_1.

The parameter δ which determines the form of the law of diminishing returns is given by

$$\delta = \frac{bc}{ad} = \frac{\left[\frac{1}{\mu_o^2} + \sigma_o^2 + 2(\tau-\tau^*)\right]\left[\frac{1}{\mu_1^2} + \sigma_1^2 + \tau^*\right]}{\tau^*\cdot\left[\frac{1}{\mu_o\mu_1} + \frac{\sigma_o^2}{\mu_1} + \frac{1}{\mu_o^2\mu_1} + \frac{\sigma_1^2}{\mu_o} + \frac{\tau^*}{\mu_o\mu_1}\right]}$$

where $\quad \tau^* = 2\left(\tau - \frac{1}{\mu_o} - \frac{1}{\mu_1}\right) \qquad (26)$

measures the tightness of the delay constraint. A straightforward calculation yields

$$\delta = 1 - \frac{\left(\frac{1}{\mu_o^2} + \sigma_o^2\right)\left(\frac{1}{\mu_1^2} + \sigma_1^2\right)}{\left(\frac{1}{\mu_o^2} + \sigma_o^2 + \frac{\tau^*}{\mu_o}\right)\left(\frac{1}{\mu_1^2} + \sigma_1^2 + \frac{\tau^*}{\mu_1}\right)} \tag{27}$$

Thus the tighter the delay constraint, <u>ceteris paribus</u>, the smaller is δ

$$\frac{\partial\delta}{\partial\tau^*} > 0. \tag{28}$$

Now observe that

$$\frac{\partial \lambda}{\partial \delta} = \frac{4q_o q_1}{(q_o+q_1)^2 - 4\delta q_o q_1} > 0, \tag{29}$$

relaxing the time constraint increases output.

$$\frac{\partial^2 \lambda}{\partial q_i \partial \delta} > 0, \tag{30}$$

relaxing the time constraint increases the marginal product of supervisors and case workers. Moreover, the elasticity of substitution between x_o and x_1 increases with δ and τ^*.

As τ^* is relaxed completely output assumes its upper bound

$$\lim_{\tau^* \to \infty} \lambda = \lim_{\delta \to 1} \lambda = \frac{q_o + q_1}{2} - \frac{1}{2} \quad (q_o-q_1)^2 = \text{Min}\ (q_o, q_1) \tag{31}$$

since the right-hand expression in (31) is either q_o or q_1. Thus the maximal stationary flow of cases is limited by the smaller of the two capacities of operatives and supervisors.

Introducing the new variables

$$v = \frac{\lambda}{x_1} = \text{output per supervisor}$$

$$u = \frac{x_o}{x_1} = \text{span of control}$$

then equation (20) assumes the form

$$v = \frac{c}{2a} + \frac{b}{2a} u - \quad (\frac{c}{2a} + \frac{b}{2a} u)^2 - \frac{du}{a}$$

$$v = \alpha + \beta u - \quad (\alpha + \beta u)^2 - \gamma u \tag{32}$$

$$\alpha = \frac{c}{2a} > 0 \quad \beta = \frac{b}{2a} > 0 \quad \gamma = \frac{d}{a} > 0 \quad \gamma = 4\alpha\beta$$

Equation (32) shows how output per supervisor increases with his/her span of control. This increase is subject to the law of diminishing returns, as demonstrated above.

It may be asked why case work requires an organization, why it should not be handled by individual case workers acting independently on their own? Presumably some kind of supervision is required to insure <u>conformity</u> with standards and <u>uniformity</u> in the application of criteria. Organization is here needed for <u>technical</u> rather than economic reasons.

We may elaborate this important point as follows. Certain tasks cannot be performed by individuals on their own but require team work or collaboration of several persons. An example are those jobs that require "checking" by another person. Thus in case work, in scientific research, in financial business deals. This checking could be done on a contractual basis by outside controllers, but when the volume of business is large enough it is cheaper to combine agents and controllers in an organization. The controlling function is then naturally associated with supervision.

12. A Production Function for Management

The considerations of the last chapter will now be extended to an organization with multiple levels of supervision.

The leading question is: how can the management process be described in terms of inputs and outputs? We shall make use once more of the concept of a production function familiar from the description of the production of goods and services. The special service produced in a management hierarchy is "management". It is not directly observable but may be inferred from the final product produced jointly by operatives and the inputs of management at various levels.

Consider a multi-level organization and a representative manager at level r. The supervisory effort achieved by the rth level of management is itself a product of managerial labor at that level and supervision from higher levels. At every level r of management, except the president, we observe the following schema

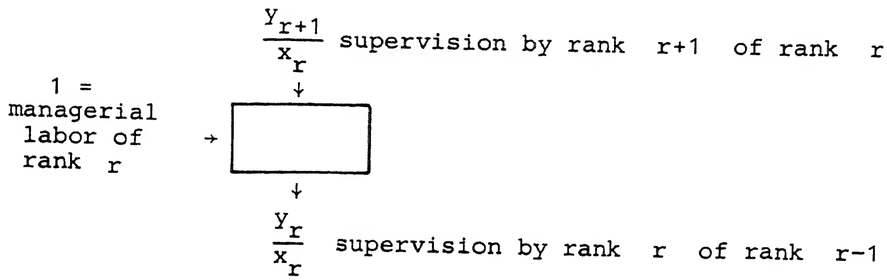


Figure 11. Schema of Managerial Control

In the black box "supervision" or managerial control is produced from two inputs managerial labor at level r and supervision by level $r+1$.

The production unit is one office. The inputs are first one unit of labor, i.e., the time of a manager of rank r, and secondly a certain amount of supervision by higher level management. The output is supervision of managers at level $r-1$.

To quantify the inputs and outputs let x_r denote the managerial labor put in at level r. "Supervision" y is both an input and an output. Total supervision y_{r+1} produced at level $r+1$ is divided among the x_r officers of rank r so that each manager of rank r receives an amount $\frac{y_{r+1}}{x_r}$ of supervision. This plus one unit of managerial labor are the inputs into one office of rank r. Total supervision produced by officers of rank r is y_r. Per office we have therefore an output $\frac{y_r}{x_r}$. The inputs into and output of a representative office of rank r is thus described by a relationship

$$\frac{y_r}{x_r} = F_r\left(1, \frac{y_{r+1}}{x_r}\right)$$

We will call

$$y_r = x_r F_r\left(1, \frac{y_{r+1}}{x_r}\right) \tag{1}$$

a "management production function".

By construction this production function is linear homogeneous. With linear homogeneous F_r an equivalent statement of (1) is

$$y_r = F_r(x_r, y_{r+1}) \qquad (1a)$$

Supervision or "management" appears here as an intermediate product which cannot be observed directly. What is observable are the final output Q by the operatives at the lowest level $r = 0$, and the labor inputs x_r at all levels of the administrative hierarchy.

Notice that Figure 11 describes the input-output relationships for a single office. The management process described here may be called "management by delegation". Each office is working on its own but subject to control or supervision by higher levels.

An alternative would be "management by team work". This will be studied in the context of a simple organization involving only one supervisor and operatives (cf. Section 15). We have indexed the production function to indicate that it may depend on the level r. If one is willing to make the assumption that "management is management" and hence is the same operation at all levels, the production function may be considered to be independent of level r.

$$y_r = F(x_r, y_{r+1}) \qquad (2)$$

At the lowest level y_o represents the organization's output

$$y_o = Q \qquad (3)$$

and at the highest, the presidential level, there is no supervisory input, the president's input is his output.

$$y_R = x_R = 1 \tag{4}$$

Successive substitution in the recursive equation (2) yields an <u>organizational production function</u>

$$Q = F_o(x_o, F_1(x_1, \ldots, F_{R-1}(x_{R-1}, x_R)) \ldots) \tag{5}$$

This is a nested production function. Since its building blocks F_r are all linear homogeneous, so is (5). When the presidential input is fixed at unity, then in terms of the remaining inputs the production function no longer exhibits constant return to scale but diminishing returns.

Suppose, for example, that supervision consists in checking and correcting the work of subordinates. The probability of detecting a hidden object through random search considered as a function of search effort (in time units) t is

$$1 - e^{-\delta t} \tag{6}$$

where $\delta\Delta t$ is the detection probability during a small time interval Δt.

Suppose further that with probability q work of a subordinate contains a serious flaw, reducing the value of the subordinate's output by a factor a, $0 < a < 1$. The expected value of the subordinate's output without supervision is then

$$1 - q + q(1-a) = 1 - qa \tag{7}$$

With supervision q is changed to $qe^{-\delta t}$.

Now the supervisory effort $\frac{y_{r+1}}{x_r}$ applied to every employee of rank r is the effective "search effort" so that the expected value of an employee's output becomes

$$\frac{y_r}{x_r} = 1 - qae^{-\delta \frac{y_{r+1}}{x_r}}$$

which is a monotone increasing and concave function of its argument. The managerial production function is then

$$y_r = x_r \, [1 - qae^{-\delta \frac{y_{r+1}}{x_r}}] \tag{8}$$

and this is independent of r provided the search effectiveness δ, the error probability q and the loss rate a are independent of r.

The special case of a Cobb-Douglas function has been studied in [Beckmann 1977]. The managerial production function has the form

$$\frac{y_r}{x_r} = a(r) \, (\frac{y_{r+1}}{x_r})^{\beta} \qquad r = 1, \ldots, R-1 \tag{9}$$

Notice that the output elasticities are assumed to be the same at all levels, but that the productivities $a(r)$ may vary. Equation (9) may be rewritten

$$y_r = a(r) \, x_r^{\alpha} \, y_{r-1}{}^{\beta} \qquad \alpha + \beta = 1 \tag{10}$$

Nesting leads to the following organizational production function

$$Q = a(0)x_o^{\alpha} \, [a(1)x_1^{\alpha}[\cdots \, a(R-1)x_{R-1}{}^{\alpha} \;\; x_R^{\beta}]^{\beta} \ldots]^{\beta}$$

$$= a_R \prod_{r=0}^{R} x_r^{\alpha_r}$$

with

$$a_R = \prod_{r=0}^{R-1} a(r)^{\beta^r}$$

$$\alpha_r = \alpha\beta^r \qquad r = 0, \ldots, R-1$$

$$\alpha_R = \beta^R$$

With the simplifying assumption that the managerial Cobb-Douglas functions are identical and the nested the organizational production function becomes

$$Q = a^{1+\beta+\ldots+\beta^{R-1}} \prod_{r=0}^{R-1} x_r^{\alpha\beta^r} \cdot x_R^{\beta^r} \tag{11}$$

In view of $x_R = 1$ the last term $x_R^{\beta^R}$ may be dropped. The organizational production function may now be used to study the allocation of management resources in an organization.

13. Allocation of Inputs: Cost Functions

In Chapter II the notion of efficiency has been introduced and refined step by step

i. The minimal number of supervisors required to supervise the execution of a given task hence the minimum number of personnel in the organization.

ii. Avoidance of slack.

iii. The minimum height of an organization consistent with a given task.

iv. the minimum average rank, minimum average salary cost, and minimum unit labor cost.

These definitions were based on given constant spans of control. In this section efficiency will be redefined as

Minimizing total salary cost incurred in achieving a given task by choosing appropriate spans of control at each level.

As mentioned before other costs (rent of office space, capital cost) could be included but this is not done here.

An alternative but equivalent efficiency concept is that for a given budget B, the organization's output is maximized.

Since we are not concerned with the external problems of market structure, we assume perfect competition in the markets for labor of all administrative ranks r.

An important assumption is that salary of staff depends only on the administrative level r. Thus we ignore salary increments due to length of service. Alternatively, we may consider a "representative" holder of a position or rank r whose salary may

contain an average component due to length of service. But the principal part of salaries must be assumed to be determined by rank, in agreement with the analysis of Section 9.

13.1 Short Run. In the short run, only labor input at the operative level $r = 0$ can be changed. The production function is then, in effect,

$$Q = F(x_o, \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_R) \tag{1}$$

with $$x_r = \bar{x}_r \tag{2}$$

held constant for $r > 0$.
Inverting the monotone function $F(x_o, \ldots)$ yields

$$x_o = \phi(Q, \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_R) \quad \text{(say)}. \tag{3}$$

As an example consider the case work production function

$$\frac{1}{\mu_o - \frac{Q}{x_o}} + \frac{1}{\mu_1 - \frac{Q}{\bar{x}_1}} = \tau \tag{4}$$

Its inverse is

$$x_o = \cfrac{1}{\mu_o - \cfrac{1}{\tau - \cfrac{1}{\mu_1 - \cfrac{Q}{\bar{x}_1}}}} \tag{5}$$

This function is defined for $Q < \mu\bar{x}_1$ and is monotone increasing and convex.

Consider also a Cobb-Douglas production function

$$Q = a_R x_o^{\alpha_o} (\bar{x}_1)^{\alpha_1} (\bar{x}_2)^{\alpha_2} \dots (\bar{x}_R)^{\alpha_R} \qquad \sum_{r=0}^{R} \alpha_r = 1 \quad (6)$$

$$= a_R x_o^{\alpha_o} \qquad \text{(say)}.$$

The inverse is

$$x_o = [a_R^{-1} \quad Q]^{\frac{1}{\delta}} \tag{7}$$

Now we introduce the cost function

$$C = \sum_{r=0}^{R} w_r x_r \tag{8}$$

Observing (2), (3) this becomes

$$C = \sum_{r=1}^{R} w_r \bar{x}_r + w_o \quad \phi(Q, \bar{x}_1, \dots, \bar{x}_R)$$

$$= F + w_o x_o(Q) \qquad \text{(say)}$$

Thus in the Cobb-Douglas case

$$C = F + w_o \, [\bar{b}_R^{-1}]^{\frac{1}{\alpha_o}} \, Q^{\frac{1}{\alpha_o}} \tag{9}$$

Observe that the variable cost has constant elasticity. The elasticity coefficient is in fact $\frac{1}{\alpha_o}$, the inverse of the output

elasticity α_o of operatives. In the case of a two-level case working organization with $\bar{x}_1 = 1$ the cost functions becomes

$$C = w_1 + \cfrac{w_o Q}{\mu_o - \cfrac{1}{\tau - \cfrac{1}{\mu_1 - Q}}}$$

This cost function is graphed in Figure 12 for $\mu_o = \mu_1 = 3$, $\tau = 1$, $w_o = 1$, $w_1 = 1.5$

$$C = 1.5 + \cfrac{Q}{3 - \cfrac{1}{1 - \cfrac{1}{3-Q}}}$$

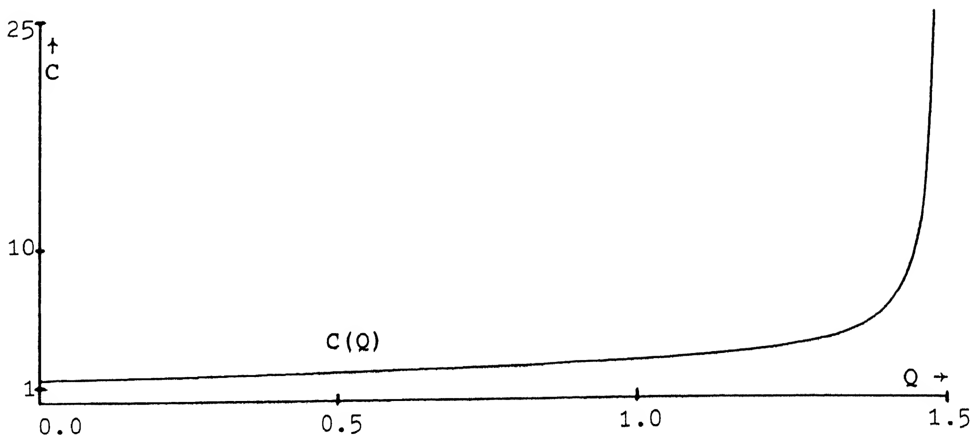


Figure 12

A Short Run Cost Function

13.2 Medium Run. In the medium run the organization can adjust personnel x_r at all levels r except the president's $r = 0, \ldots, R-1$, $\bar{x}_R = 1$.

The organization's production function is, therefore,

$$Q = F(x_o, x_1, \ldots, x_{R-1}, 1)$$

Thus in the case working case Q is given implicitly by

$$\tau = \sum_{r=0}^{R} \frac{\Pi_r}{\mu_r - \frac{Q\Pi_r}{x_r}} \qquad x_R = 1 \qquad (10)$$

where Π_r is the probability of a case getting to level r. In the Cobb-Douglas case

$$Q = a_R \cdot x_o^{\alpha_o} \; x_1^{\alpha_1} \; \ldots \; x_{R-1}^{\alpha_{R-1}} \tag{11}$$

Obtaining the cost minimizing or efficient combinations of personnel x_r in the medium run in order to handle a given task Q is tantamount to finding the optimal spans of control

$$x_{r-1}/x_r, \qquad r = 1, \ldots, R.$$

Mathematically the problem is to determine

$$\text{Min} \sum_{r=0}^{R-1} w_r x_r + w_R$$

subject to production functions (10) or (11). In the case-working case one obtains

$$x_r = \frac{\Pi_r}{\mu_r} [Q + \sqrt{\psi Q/w_r}] \qquad r = 0, \ldots, R-1 \tag{12}$$

where the Lagrangean multiplier ψ is determined by (10)

$$\tau = \sum \frac{\Pi_r}{\mu_r} \frac{1}{1 - \frac{Q}{Q + \sqrt{\psi Q/w_r}}} \tag{13}$$

From (13) it is seen that ψ is a decreasing function of τ, the delay time. In the following we may use the Lagrangian multiplier ψ as a quality measure of output instead of τ.

The probability of a case reaching level r is a decreasing function of r. Since the difficulty of these cases increases, the rate of disposal μ_r presumably decreases with r. It is not unreasonable that the ratio $\frac{\Pi_r}{\mu_r}$ is approximately constant.

Since w_r increases it follows that the number x_r of employees decreases with rising rank r. The span of control for constant $\frac{\Pi_r}{\mu_r}$ is then given by

$$s_r = \frac{x_{r-1}}{x_r} = \frac{Q + \frac{\psi Q}{w_{r-1}}}{Q + \frac{\psi Q}{w_r}}$$

For a constant salary span $\frac{w_{r+1}}{w_r} = b$ one has

$$s_r = \frac{Q + \frac{\psi Q b}{w_r}}{Q + \frac{\psi Q}{w_{r+1}}}$$

and this is a decreasing function of w_r and hence of r, approaching $s_r = 1$ as $w_r \to \infty$.

In the Cobb-Douglas case we have instead a Lagrangean

$$L = - \sum_{r=0}^{R} w_r x_r + \psi \cdot [a_R \prod_{r=0}^{R-1} x_r^{\alpha_r} - Q] \tag{14}$$

yielding

$$x_r = \psi \cdot \frac{\alpha_r Q}{w_r} \tag{15}$$

Upon substituting (15) in the production function (11) one obtains

$$\psi = a_R^{-\frac{1}{1-\alpha_R}} Q^{\frac{\alpha_R}{1-\alpha_R}} \prod_{r=0}^{R-1} \left(\frac{w_r}{\alpha_r}\right)^{\frac{\alpha_r}{1-\alpha_R}} \tag{16}$$

The span of control has the simple form

$$s_r = \frac{x_{r-1}}{x_r} = \frac{\alpha_{r-1}}{\alpha_r} \cdot \frac{w_r}{w_{r-1}} \qquad r = 1,\ldots,R-1 \tag{17}$$

This shows clearly how the span of control depends on relative wage rates.

This remarkable formula may be read in a slightly different way:

$$w_r = \frac{\alpha_r}{\alpha_{r-1}} s_r w_{r-1}. \tag{18}$$

Since $s_r w_{r-1}$ is the combined salary of those supervised (on the average) by an administrator level r, (18) states that the span of control should establish a proportional relationship between salaries of supervisors and those of the supervised [cf. Lydall, 1968].

In the case of identical Cobb-Douglas management production functions equation (17) shows that

$$s_r = \frac{1}{\beta} \frac{w_r}{w_{r-1}} \tag{19}$$

Once more, when salaries increase by a constant factor b

$$s_r = s = \frac{b}{\beta} \tag{20}$$

Equation (20) states that the average span of control should be constant at all levels r. This optimal span of control is

proportional to the incremental salary factor b and inversely proportional to the output elasticity of supervision β.

Formula (16) permits to give rough estimates of α and β. If we accept as representative values

$$b = 1.25$$

$$s = 5$$

then

$$\beta = \frac{g}{\beta} = .25$$

This is in agreement with the estimated output elasticity of labor $\alpha = 75$ in general Cobb-Douglas production functions.

13.3 <u>Cost Curves</u>. Substitute now the recommended inputs (15) into the cost

$$C = \sum_{r=0}^{R} w_r x_r$$

to obtain

$$C = w_R + [1 - \alpha_R]\,\psi Q$$

where we have used

$$\sum_{r=0}^{R} \alpha_r = 1$$

With the expression (16) substituted for ψ one obtains

$$C = w_R + (1-\alpha_R)\, a_R^{-\frac{1}{1-\alpha_R}}\, Q^{\frac{1}{1-\alpha_R}} \prod_{r=0}^{R-1} \left(\frac{w_r}{\alpha_r}\right)^{\frac{\alpha_R}{1-\alpha_R}} \qquad (21)$$

Once more, there is a fixed cost w_R and a variable cost whose elasticity is constant. The value of the elasticity

$$\eta = \frac{1}{1-\alpha_R} \qquad (22)$$

is smaller than the short-run cost elasticity $\frac{1}{\alpha_o}$ since

$$\alpha_o < 1 - \alpha_R$$

in view of $\sum_{r=0}^{R} \alpha_r = 1$

Equation (22) states that the elasticity of variable cost with respect to output decreases with the number of levels R. In the special case of management production functions with identical exponents α, β one has $\alpha_R = \beta^R$ so that

$$\eta = \frac{1}{1-\beta^R}$$

which is clearly decreasing with respect to R.

Figure 13 shows some medium-run cost curves for $\alpha = \beta = \frac{1}{2}$ $a = 2.2$, $w_r = (\frac{3}{2})^r$.

13.4 Long Run. In the long run the level R of the organization may also be adjusted. The optimum level R which produces Q at minimum cost may be read off diagrams like Figure 13 or the appropriate cost curves that apply when the managerial production function is not Cobb-Douglas.

As the size of the task Q increases, so will the general number of levels R. This is best seen by comparing those outputs Q_R for which unit cost is minimized for given R.

A straightforward calculation for identical Cobb-Douglas production function yields [Beckmann, 1978, p. 128, equation (17)]

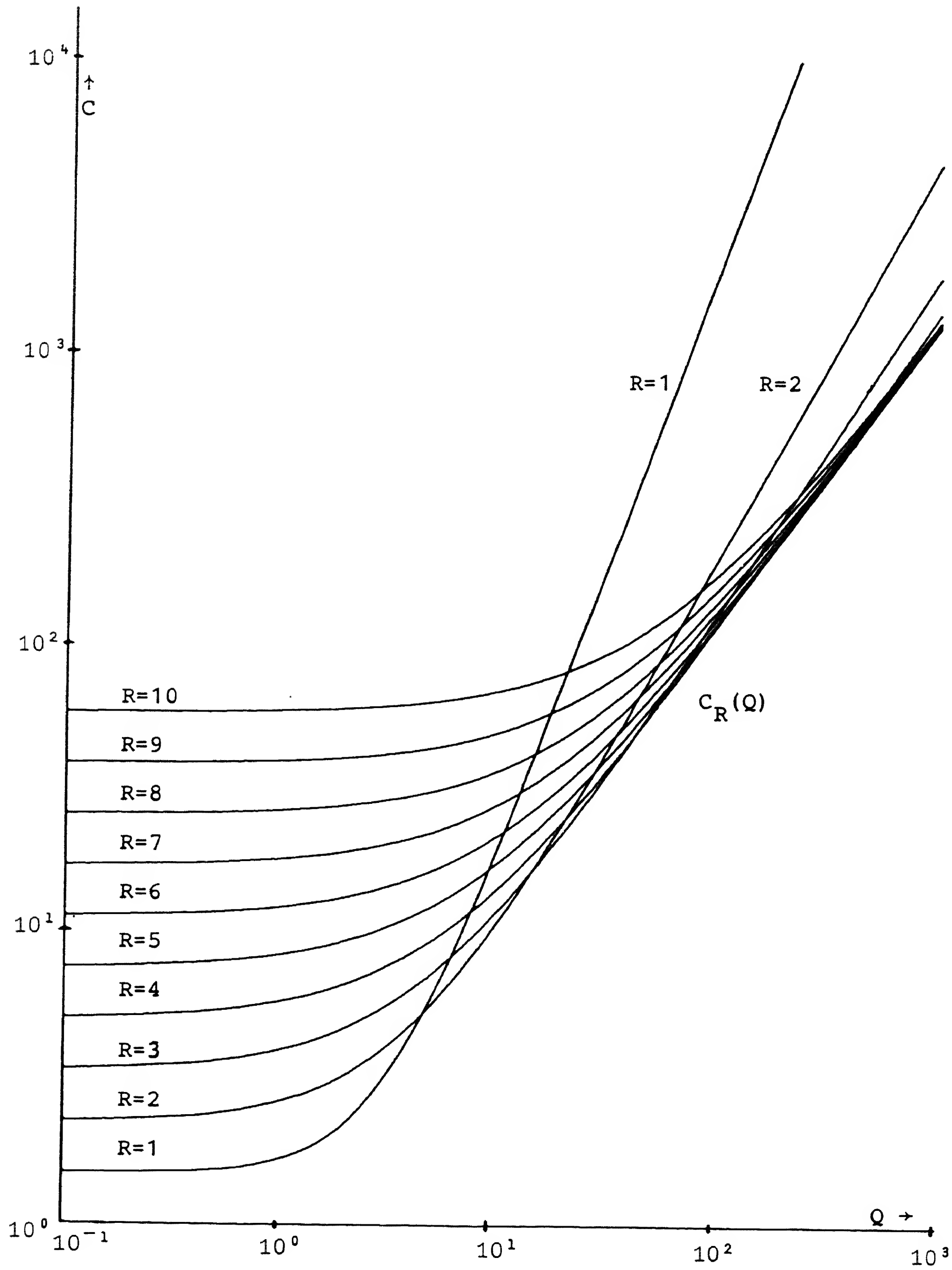


Figure 13. Cost Curves for Different R

$$Q_R = a_R s_R^{1-\beta^R} s^{R + \frac{\beta^R - 1}{\alpha}} \tag{23}$$

Here s_R and s are determined by the salary scale. In particular for constant proportional increments it follows that (for large R since $\beta^R \to 0$)

$$Q_R \doteq Q_o \cdot s^R \tag{24}$$

where s is given by (20) $s = \frac{b}{\beta}$

Thus the optimum output increases exponentially with the organization's top rank R. The top rank of R is strictly proportional to the logarithm of optimum output.

Another important question is how average cost (or total cost) varies with output. This topic will be taken up in fuller generality in Section 17.

At this point we remark the following: When the management production functions are identical Cobb-Douglas

$$F_r(x_r, y_{r+1}) = a x_r^{\alpha} y_{r+1}^{\beta} \qquad \alpha + \beta = 1$$

then minimum unit cost for an organization with R levels achieved by an output Q_R as given in (24) is

$$c_R = w_o a^{-\frac{1}{\alpha}} b^{\frac{\beta}{\alpha}} \alpha^{-1} \beta^{-\frac{\beta}{\alpha}} \cdot \left[\frac{a b_R^{\alpha}}{b} \alpha^{\alpha} \beta^{\beta}\right]^{\frac{\beta}{\alpha}} \tag{25}$$

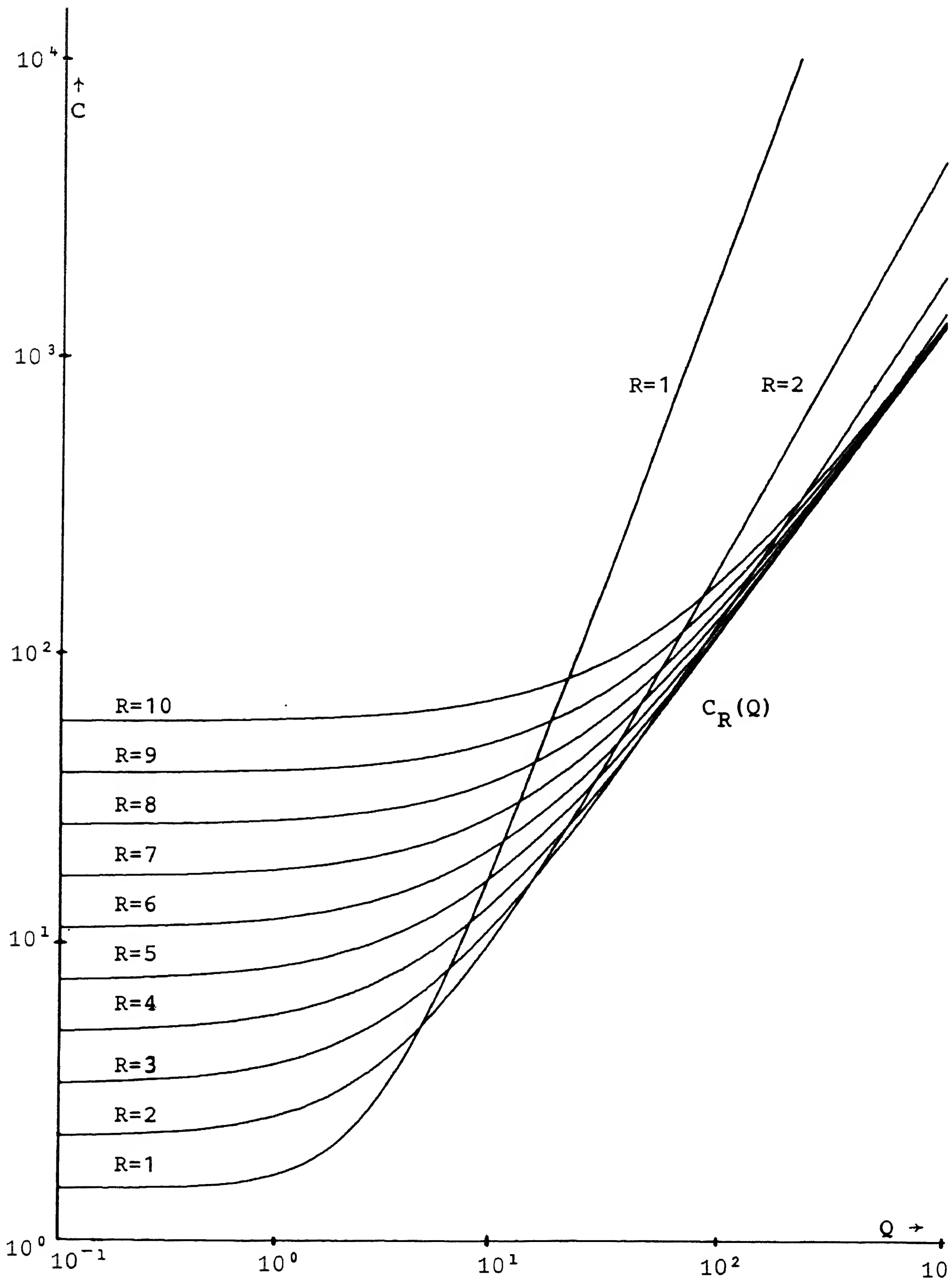


Figure 13. Cost Curves for Different R

$$Q_R = a_R s_R^{1-\beta^R} s^{R + \frac{\beta^R - 1}{\alpha}} \tag{23}$$

Here s_R and s are determined by the salary scale. In particular for constant proportional increments it follows that (for large R since $\beta^R \to 0$)

$$Q_R \doteq Q_o \cdot s^R \tag{24}$$

where s is given by (20) $s = \frac{b}{\beta}$
Thus the optimum output increases exponentially with the organization's top rank R. The top rank of R is strictly proportional to the logarithm of optimum output.

Another important question is how average cost (or total cost) varies with output. This topic will be taken up in fuller generality in Section 17.

At this point we remark the following: When the management production functions are identical Cobb-Douglas

$$F_r(x_r, y_{r+1}) = a x_r^\alpha \, y_{r+1}^\beta \qquad \alpha + \beta = 1$$

then minimum unit cost for an organization with R levels achieved by an output Q_R as given in (24) is

$$c_R = w_o a^{-\frac{1}{\alpha}} b^{\frac{\beta}{\alpha}} \alpha^{-1} \beta^{-\frac{\beta}{\alpha}} \cdot \left[\frac{a b_R^\alpha}{b} \alpha^\alpha \beta^\beta\right]^{\frac{\beta}{\alpha}} \tag{25}$$

The conclusion is that this minimum unit cost is asymptotically constant, for

$$\lim_{r\to\infty} \beta^R = 0 \qquad \text{implies}$$

$$c_R \to w_o a^{-\frac{1}{\alpha}} b^{\frac{\beta}{\alpha}} \alpha^{-1} \beta^{-\frac{\beta}{\alpha}} \tag{26}$$

V. Advantage of Organization

V. ADVANTAGE OF ORGANIZATION

14. Organization versus the Individual.

In Section 11 and elsewhere we have considered tasks that can only be performed by organizations. There are other activities in which organizations compete with individuals. In many areas it appears that organizations are in fact on the advance. Increasingly they are taking over functions previously performed mainly by individuals. Is this true? What, if any, are the economic facts that would permit and drive this process?

An economist will ask more specifically: What is the economic advantage, if any, that an organization has in doing jobs otherwise performed by indivdiuals? This is the question studied in the present section. The following Section 15 compares two basic types of organization: partnerships and organizations with subordination and supervision. Section 16 examines how multi-level organization or hierarchies proper can enter this competition. In Supplement H we examine how personnel of different qualifications should be utilized by hierarchical organizations.

One assumption made throughout is that some operatives are also qualified to be supervisors or managers. One may even assume that everybody is equally qualified to be a manager (although this is neither realistic nor necessary). In that case one would add an assumption that management jobs require more effort than operative labor. Casual empiricism suggests that those who found organizations and manage them (entrepreneurs) put in longer working hours than do operatives.

When comparing the output of organizations with that of individuals on their own, as e.g., in the professions, one must address the question of how input increases as the result of "organization". Now organization has at least two aspects: specialization and supervision or "management". The need for management decreases the number of person units available for operative labor in an organization of n members. At the same time specialization should increase their productivity. The output of the organization as measured by an organizational production function should reflect both factors. The analysis of this section is in terms of simple organizations defined as involving only one level of supervision.

A person operating on his/her own is capable of producing (say) one unit of output per unit of time, valued at unity. Unity is, therefore, the opportunity wage of qualified personnel.

Even an individual on his/her own must allocate, however, some time to management functions, e.g., to the planning of work, to contacting customers, to decisions about capital equipment, etc. Thus if an amount x_o is allocated to operative work and x_1 to management, the result will be

$$F(x_o, x_1).$$

The output y attainable with a total time input x is then given by

$$y = \max_{x_o + x_1 \leqq x} F(x_o, x_1)$$

or

$$y = \phi(x), \qquad \text{say.} \tag{1}$$

By definition we have

$$\phi(0) = 0 \qquad \phi(1) = 1. \tag{2}$$

Proposition 1. Let $F(x_o,x_1)$ be homogeneous of degree m. Then

$$\phi(x) = x^m \quad \text{for all} \quad x \leqq 1$$

Proof. $(\frac{1}{x})^m F(x_o,x_1) = F(\frac{x_o}{x}, \frac{x_1}{x})$

$$\phi(x) = \underset{x_o+x_1=x}{\text{Max}} F(x_o,x_1)$$

$$= x^m \underset{\frac{x_o}{x} + \frac{x_1}{x} = 1}{\text{Max}_{x_o\, x_1}} F(\frac{x_o}{x}, \frac{x_1}{x})$$

Writing $\frac{x_o}{x} = u \quad \frac{x_1}{x} = v$

$$\phi(x) = x^m \underset{\substack{u,v \\ u+v=1}}{\text{Max}} F(u,v)$$

$$= x^m \phi(1) = x^m \qquad \text{Q.E.D.}$$

Therefore, if the combination of management and operative labor in one individual has returns to scale of constant degree $m > 0$, the resulting production function ϕ is a power function with exponent m. Consider the case of decreasing returns to scale $m < 1$. Suppose a unit of work is divided among $\frac{1}{x}$ persons each performing an amount of work $x < 1$. The result is a total output exceeding unity since $\frac{1}{x} \cdot x^m = x^{m-1} > 1$ for $x < 1$, and in fact the finer the division, the greater the total result. This is an empirically unlikely event.

In the following we consider the alternative case $m > 1$ constant or increasing returns.

Consider now output exceeding unity, i.e., the capacity of one individual. A simple organization is now required. The question is how much person time shall be allocated to management.

In small organizations, a person designated as manager, whether permanently or on a rotating basis, may still want to devote part of his/her time to operative labor. But no more than one person may perform management functions in a simple organization with only one supervisory level. The production function for this type of organization is therefore

$$\phi(x) = \max_{\substack{x_0+x_{10}+x_{11} \leqq x \\ x_{11} \leqq 1}} F(x_0 + x_{10}, x_{11})$$

Proposition 2.

Let $F(u,v) = 1 = \max_{\substack{x_0, x_1 \\ x_0+x_1=1}} F(x_0, x_1)$

u,v are the output maximizing allocations of labor to operative work and management in producing a unit output. Then

$$\phi(x) = x^m$$

for all $x \leqq \frac{1}{v}$.

Proof. The constraint $x_{11} \leqq 1$ is inoperative as long as $vx \leqq 1$. Therefore, the previous production functions $\phi(x)$ still applies to all $x \leqq \frac{1}{v}$. Q.E.D.

All larger outputs are produced under the additional constraint $x_{11} = 1$ which implies $x_{10} = 0$.

$$\phi(x) = \max_{\substack{x_0+x_1=x \\ x_1 \leqq 1}} F(x_0, x_1)$$

$$= \max_{\substack{x_0+x_1=x \\ x_1=1}} F(x_0 x_1)$$

or

$$\phi(x) = F(x-1,1) \qquad x \geqq \frac{1}{v}. \tag{4}$$

Varying one input (operative labor) while holding the other constant (management) is the classical situation to which the law of diminishing returns to substitution applies. We postulate

Assumption. $F(x_0,x_1)$ for x_1 = constant is a concave function of x_0.

This implies immediately

Proposition 3. $\phi(x)$ is a concave function of x for $x \geqq \frac{1}{v}$.

The production function ϕ has, therefore, the following characteristics for $m \geqq 1$.

$$\phi(x) = \begin{array}{lll} x^m, & \text{convex} & \text{for } x \leqq \frac{1}{v} \\ F(x-1,1), & \text{concave} & \text{for } x \geqq \frac{1}{v}. \end{array} \tag{5}$$

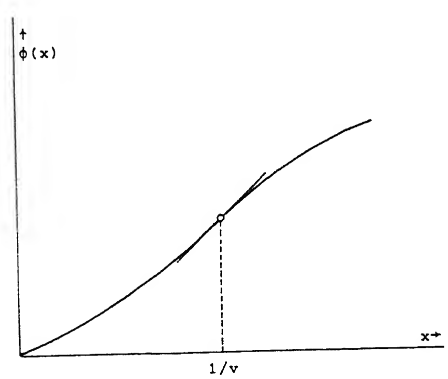


Figure 14. Output of a Simple Organization

As an illustration, consider a Cobb-Douglas function

$$F(x_0 x_1) = a x_0^{\alpha} x_1^{\beta} \qquad \alpha + \beta = m \geqq 1 \tag{6}$$

$$\phi(x) = \max_{x_0+x_1=x} a x_0^{\alpha} x_1^{\beta}$$

implies

$$x_0 = \frac{\alpha}{\alpha+\beta} x \qquad \text{or} \quad u = \frac{\alpha}{\alpha+\beta}$$

$$x_1 = \frac{\beta}{\alpha+\beta} x \qquad\qquad v = \alpha^{\beta}{+}\beta$$

$$\phi(x) = a\alpha^{\alpha} \beta^{\beta} (\alpha+\beta)^{-\alpha-\beta} x^{\alpha+\beta}.$$

From (2), we have

$$m = \alpha+\beta$$

$$a = \alpha^{-\alpha} \beta^{-\beta} (\alpha+\beta)^{\alpha+\beta}$$

$$\phi(x) = \begin{cases} x^{\alpha+\beta} & x \leqq \dfrac{\alpha+\beta}{\beta} \\ (\alpha+\beta)^{\alpha+\beta} \alpha^{-\alpha} \beta^{-\beta} (x-1)^{\alpha} & x \geqq \dfrac{\alpha+\beta}{\beta} \end{cases} \tag{7}$$

Observe that $\phi(x)$ is continuously differentiable at $x = \frac{1}{v}$, but that its second derivative is, in general, discontinuous there.

Is organization advantageous? We show:

Proposition 4. Suppose F is homogeneous of degree m. Then the return to organization k exceeds or equals the opportunity wage of unity according as m is greater or equal to one.

$$k \left\{ \substack{>\\=} \right\} 1 \iff m \left\{ \substack{>\\=} \right\} 1 \tag{8}$$

Proof: A maximum of

$$\phi(x) - (x-1)$$

occurs only where ϕ is concave, hence for $x > \frac{1}{v}$. Then

$$\phi(x) = F(x-1,1).$$

Now

$$F(x-1,1) = x^m F(\frac{x-1}{x}, \frac{1}{x})$$

$$= x^m \quad \text{for} \quad x = \frac{1}{v}$$

so that

$$k = \underset{x}{\text{Max}} \quad F(x-1,1) - (x-1) \geqq (\frac{1}{v})^m - (\frac{1}{v} - 1)$$

$$= 1 + \frac{1}{v} (v^{1-m}-1) > 1 \qquad (9)$$

since $v < 1$, $m > 1$.

If $m = 1$ then

$$\phi(x) \leqq x \qquad \text{and}$$

$$= x \quad \text{for} \quad x \leqq \frac{1}{v}$$

from which

$$k = \underset{x}{\text{Max}} \quad \phi(x) - (x-1) = 1 \qquad (9)$$

We conclude that under constant returns to scale the reward for organization is just equal to the wage. Therefore, organizations cannot do any better than indivdiuals. There is no economic reason for organization.

If $m < 1$ then

$$\phi(x) = x^m \qquad x \leqq \frac{1}{v}$$

is concave and

$$\phi(x) - x$$

which vanishes for $x = 0$ and $x = 1$ has a maximum when

$$x = m^{\frac{1}{1-m}} < 1 \qquad \text{yielding}$$

$$\max_x \; x^m - x = m^{\frac{m}{1-m}} - m^{\frac{1}{1-m}} = (1-m)m^{\frac{1}{1-m}} > 0 \qquad (10)$$

as a surplus achieved with $x < 1$, i.e., through part-time work. Under competition with free entry, this surplus must vanish and consequently the returns to full-time work are less then unity: only part-time work will pay the opportunity wage of unity. Such cases of optimum performance under part-time work exist notably in the domestic area, but they are not of interest in the study of organizations.

15. Supervision vs. Partnerships

Why hierarchical organizations? This question is attracting attention among economists [Williamson, 1975, Beckmann, 1978, Calvo, 1979, Rosen, 1981]. A complete analysis must draw on sources in various disciplines, but clearly involve economic reasoning as well. In this section, we address a slightly narrower question: Under what conditions is a hierarchical organization preferable to the alternative of a partnership? We assume that technically the organization can operate at any level. It is conceivable that the organization's output can be produced also by individuals operating on their own, but this need not be assumed here.

15.1 An organization whether hierarchical or partnership is distinguished from individual activity by a division of labor. Some of its manpower must be engaged in management, some in operative labor. We denote the output of the management and operative effort by a production function

$$F(x_o, x_1) \tag{1}$$

where y output

x_0 operative effort

x_1 management effort.

Management is assumed to be indivisible, i.e., it cannot be shared--but it may be rotated. If each person's capacity (or limit of effort) is unity, this means

$$x_1 \leqq 1. \tag{2}$$

As before, we assume that both the value of a unit of output and the going wage rate for operative labor are unity. Also persons producing on their own produce just one unit of output so that unity is the opportunity wage. A hierarchy and a partnership are then distinguished first of all by the distribution of the income generated by the organization's output.

Assume that a partnership shares the income in proportion to each person's level of effort. In particular, if each person puts in one full-time equivalent of effort, this means an equal share for everybody.

$$y_p = \frac{F(x_0,x_1)}{x_0+x_1}.$$

In the special case that the production function is linearly homogeneous, each partner's income per unit of effort becomes independent of the scale of operation. The constraint (2) may be ignored. For comparisons with a hierarchical organization, we can set $x_1 = 1$, without affecting the income level of any partner.

In a hierarchical organization the manager does only manage. His income is the value remaining of output after paying wages to hired operators

$$y_h = F(x_0,x_1) - x_0.$$

We wish to compare incomes and spans of control for a partnership and a hierarchical organization.

15.2 For a partnership write

$$g = \underset{\substack{x_0, x_1 \\ x_1 \leqq 1}}{\text{Max}} \frac{F(x_0, x_1)}{x_0 + x_1}$$

$$g = \underset{x_0}{\text{Max}} \frac{F(x_0, 1)}{x_0 + 1} \quad \text{by linear homogeneity of } F.$$

$$g = \frac{F(\hat{x}, 1)}{\hat{x} + 1} \quad \text{(say).} \tag{3}$$

For a hierarchical organization write

$$k = \underset{\substack{x_0, x_1 \\ x_1 \leqq 1}}{\text{Max}} \; F(x_0, x_1) - x_0$$

$$k = \underset{x_0}{\text{Max}} \; F(x_0, 1) - x_0$$

since the production function F is monotone increasing in both inputs.

$$k = F(x^*, 1) - x^* \quad \text{(say)} \tag{4}$$

We now introduce the assumption that a partnership is profitable, i.e., can achieve more per person than an individual on his/her own

$$g > 1.$$

Now from (3)

$$
\begin{aligned}
g &= F(\hat{x},1) - g\hat{x} \\
&< F(\hat{x},1) - x \\
&\leqq \max_x \; F(x,1) - x \\
&= k \qquad\qquad \text{by (4).}
\end{aligned}
$$

Therefore,

$$g < k. \tag{5}$$

We conclude: If a partnership is profitable, a hierarchical organization is even more profitable for the organizer.

A third possibility is that of workers hiring a manager at a salary w between what owners demand (h) and what partners receive (g) as their share in a partnership

$$g < w < k. \tag{6}$$

Now by definition of g

$$\frac{F(x,1)}{x+1} \leqq g$$

$$F \leqq gx + g$$

$$\frac{F-g}{x} \leqq g.$$

Therefore,

$$\frac{F-w}{x} < g \qquad\qquad \text{using (6).}$$

When workers hire managers, their share falls below that of a pure partnership [cf. Putterman, 1981b].

In a partnership, the optimal ratio $\frac{x_0}{x_1} = \frac{\hat{x}}{1}$ is independent of the level of operation provided $x_1 \leqq 1$. Suppose the production function is continuously differentiable and concave. Then $\hat{x}$ is determined by

$$0 = \frac{\partial}{\partial x} \frac{F(x,1)}{x+1} = \frac{F_1(\hat{x},1)\cdot[\hat{x}+1] - F(\hat{x},1)}{(\hat{x}+1)^2}$$

or

$$F_1(\hat{x},1) = \frac{F(\hat{x},1)}{\hat{x}+1} = g > 1. \qquad (6)$$

The optimal span of control in a hierarchical organization is similarly given by

$$0 = \frac{\partial}{\partial x} [F(x,1)-x] = F_1(x^*,1) - 1$$

from which

$$F_1(x^*,1) = 1. \qquad (7)$$

Concavity of F implies

$$F_{11}(x,1) < 0. \qquad (8)$$

Combining (6), (7) and (8) one sees that

$$\hat{x} < x^*. \qquad (9)$$

The optimum span of control in a hierarchical organization is larger than that in a partnership [cf. Figure 15]. This has been found to apply to "worker managed enterprises" as compared to conventional firms by Ward [1958].

15.3 Return to the production function (1) and constraint (2) on management. What is the optimal allocation of a given number of personnel or total level of effort x to operate labor x_0 and management x_1 to achieve a maximum level of output z? This question is relevant to both partnership and to hierarchical organizations in which the boss is willing to do some operative labor

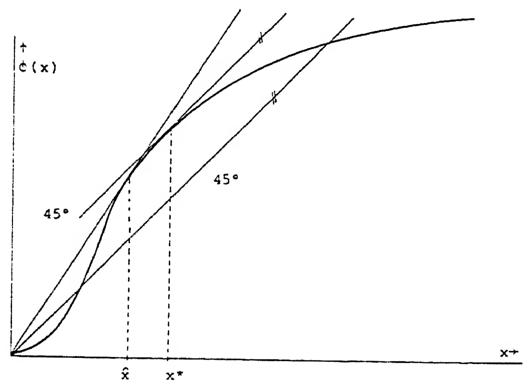


Figure 15. Optimal Control Spans in Alternative Organizations

$$z = \max_{\substack{x_0, x_1 \\ x_1 \leqq 1 \\ x_0 + x_1 \leqq x}} F(x_0, x_1) = \phi(x) \quad \text{(say)}. \qquad (10)$$

<u>Proposition</u>: If $F(x_0,x_1)$ is linear homogeneous, then at low levels of x, $\phi(x)$ is a proportional function $\phi(x) = gY$.

<u>Proof</u>: If Y is small, e.g., $x < 1$ then constraint (2) is automatically satisfied so that (10) reduces to

$$z = \max_{x_0 + x_1 = x} F(x_0, x_1)$$

which is solved by

$$F_1(x_0,x_1) = F_2(x_0,x_1). \tag{11}$$

Condition (11) determines uniquely the optimizing ratio for partnership

$$\frac{x_0}{x_1} = s = \text{optimal span of control} \tag{12}$$

which may be seen as follows. Write

$$F(x_0,x_1) = x_1 \cdot F(\frac{x_0}{x_1},1)$$

using linear homogeneity. Condition (11) becomes

$$F_1(\frac{x_0}{x_1},1) = F(\frac{x_0}{x_1},1) - \frac{x_0}{x_1} F_1(\frac{x_0}{x_1},1)$$

or

$$F(s,1) - (1+s) F_1(s,1) = 0. \tag{13}$$

But the left-hand side of (13) is an increasing function of s when F is concave (as seen by differentiating with respect to s) proving that s, if it exists, is uniquely determined. For s to exist, it is sufficient that

$$F(0,1) = 0$$

$$F_1(1,0 > 0$$

$$F(s,1) \quad \text{is bounded.}$$

Condition (12) implies

$$x_0 = \frac{s}{s+1}\, x \tag{14}$$

$$x_1 = \frac{1}{s+1}\, x \tag{15}$$

$$\phi(x) = F_1(s,1)\cdot x \qquad \text{or}$$

$$\phi(x) = g\cdot x. \tag{16}$$

Equations (14)-(16) are valid for $x_1 = \frac{1}{s+1}\, x \leqq 1$ or

$$x \leqq s+1. \tag{17}$$

This proves the assertion.

In fact, we have the stronger

Corollary: Let $F(x_0, x_1)$ be homogeneous of degree m. Then for s determined by (13) and all

$$x \leqq 1+s \tag{18}$$

$$\phi(x) = gx^m. \tag{19}$$

The proof follows along the same lines as that of the proposition. It differs from Proposition 1 of Section 14 only by the fact, that we no longer assume $\phi(1) = 1$.

To illustrate consider a linear homogeneous Cobb-Douglas production function

$$z = ax_0^\alpha\, x_1^\beta \qquad \alpha + \beta = 1 \tag{20}$$

$$k = [ab\alpha^\alpha\beta^\beta]^{\frac{1}{\beta}} \qquad g = b\alpha^\alpha\beta^\beta \tag{21}$$

$$k > g > 1 \quad \text{iff} \quad a > \alpha^{-\alpha}\ \beta^{-\beta} = 1.7545$$

for the conventional values $\alpha = \frac{3}{4}$ $\beta = \frac{1}{4}$.

It follows that

$$\hat{x} = s = \frac{\alpha}{\beta} \tag{22}$$

$$x^* = (\alpha a)^{\frac{1}{\beta}} > (\alpha\alpha^{-\alpha}\ \beta^{-\beta})^{\frac{1}{\beta}} = \frac{\alpha}{\beta} = s = 3. \tag{23}$$

15.4 Why are partnerships ever chosen if they are less profitable? Presumably because they afford other advantages such as greater freedom; for instance, the freedom to choose one's level of working time or effort.

We discuss this in terms of a specific utility function. Let t be working time, $c-t$ leisure and y income. The utility function to be considered has the simple form

$$u = \ln (c-t) + d \ln y. \tag{24}$$

We consider three alternatives.

Case 1: Each partner i receives a fixed share of profits a_i regardless of effort t_i. Subsequently we shall specify this to mean equal shares. (The "commune" of Sen [1966].)

Case 2: Reward y_i is proportional to effort, (the "collective" of Sen [1966].) Thus for partner i

$$y_i = \frac{t_i}{t} \cdot y \quad \text{where} \quad t = \sum_i t_i. \tag{25}$$

In the case of hired employees earning a unit wage for a unit of effort, utility equals

$$u = \ln (c-t_i) + d \ln t_i.$$

In the case $c = 2$, $d = 1$, this is maximized by $t_i = 1$ yielding $y = 1$ and $u = 0$ as bench marks.

Since the partnership operates in the range $x < s+1$ is linear homogeneous production function implies

$$y = gt.$$

In the fixed share case, a partner's income equals

$$y_i = g \cdot a_i \cdot (t_i + T)$$

where

$$T = \sum_{j \neq i} t_j \quad \text{is time input by the other partners.}$$

Now

$$u_i = \ln (c - t_i) + d \ln (g\ a_i (t_i + T)).$$

Maximization with respect to t_i yields

$$-\frac{1}{c-t_i} + \frac{d}{t_i+T} = 0$$

$$t_i = \frac{dc-T}{d+1}$$

$$t_i + T = \frac{d}{d+1} \cdot (T+1).$$

Notice that a partner's effort t_i is independent of a partner's share a_i. To determine T observe that, since everybody puts in the same amount of time.

$$T = (n-1)t_i$$

yielding

$$T = \frac{(n-1)d}{n+1}\ c$$

$$t_i = \frac{dc}{n+d}$$

$$T + t_i = \frac{nd}{n+d}\ c.$$

For the conventional values
$d = 1, \quad c = 2$

$$T+t_i = \frac{2n}{n+1} < 2 \leq s+1$$

for any integer span of control s.

Achieved utility is

$$\hat{u} = \ln\left(\frac{nc}{n+d}\right) + d \ln\left(a_i \cdot g \cdot \frac{nd}{n+d} c\right)$$

$$= (1+d) \ln \frac{nc}{n+d} + d \ln a_i gd.$$

In the case of equal shares

$$\hat{u} = (1+d) \ln \frac{nc}{n+d} a_i + d \ln \frac{gd}{n}. \qquad (26)$$

The optimal number of partners n is then the minimal number, since

$$\frac{\partial \hat{u}}{\partial n} = \frac{1}{n} - \frac{1+d}{n+d} < 0 \quad \text{for} \quad n > 1$$

and the minimal integral number of partners is obviously $n = 2$.

Thus small partnerships are more advantageous than larger ones when each partner receives a fixed share regardless of effort.

Utility is positive and being a partner is preferable to working as a zero level employee of a hierarchical organization when

$$(1+d) \ln \frac{nc}{n+d} + d \ln a_i gd > \ln (c-1)$$

which is true when share a_i is large enough and the number of partners n is small enough. When $c = 2$, $d = 1$, the condition is

$$\frac{(n+d)^2}{4n} < g$$

which imposes an upper limit on n. Thus

$$g = 2$$

implies $\quad n < 5.$

15.6 Turn next to the case of rewards proportional to input. Utility to the ith partner is

$$u_i = \ln (c-t_i) + d \ln (\frac{t_i}{t_i+T} \cdot g \cdot (t_i+T)$$

or

$$u_i = \ln (c-t_i) + d \ln (gt_i).$$

This is maximized by

$$t_i = \frac{cd}{d+1}.$$

Achieved utility is then

$$\tilde{u} = (1+d) \ln \frac{c}{1+d} + d \ln gd.$$

Comparison with the previous case shows that

$$\tilde{u} = (1+d) \ln \frac{c}{1+d} + d \ln gd$$

$$> (1+d) \ln \frac{nc}{1+d} a_i + d \ln a_i gd = \hat{u}$$

when $\quad a_i < \frac{1}{n}$

so that a partnership with proportional rewards yields greater utility, a result that should cause no surprise, since the fixed share system discourages effort.

This conclusion is actually true for arbitrary concave utility functions in terms of leisure and income, when shares are equal.

$$\hat{u} = \underset{t}{\text{Max}} \quad u(c-t, \frac{1}{n} f(T+t)$$

$$= u(c-\hat{t}, \frac{1}{n} f(T+\hat{t}))$$

$$\hat{u} = u(c-\hat{t}, \frac{1}{n} f(n\hat{t}))$$

since $\quad T = (n-1)\hat{t}.$

Rewrite $\quad \hat{u} = u(c-\hat{t}, \frac{\hat{t}}{\hat{t}n} f(\hat{t}n))$

$$\hat{u} \leqq \underset{t,L}{\text{Max}} \quad u(c-\hat{t}, \frac{t}{L} f(L)) = \tilde{u}.$$

Moreover, a partnership with the optimal number of members yields the utility that a benevolent dictator would achieve through central planning for the partnership (ignoring indivisibilities). For central planning would exactly yield the same result as shown by the following

Theorem:

When the partnership has the optimal number of members then the distribution proportional to effort achieves the optimum for every member provided the members' preterences are equal.

Proof:

The optimum is given by

$$\underset{t}{\text{Max}} \quad u(c-t, \frac{1}{n} \phi(nt))$$

and this is achieved by

$$0 = \frac{\partial u}{\partial t} = - u_1 + u_2 \phi'$$

$$0 = \frac{\partial u}{\partial n} = u_2 \cdot [n\phi' - \phi]/n^2$$

or $$\phi' = \frac{\phi}{nt} \quad \text{average product = marginal product.} \qquad (27)$$

Now consider the utility of an individual member under the proportional distribution.
Let T represent the time input of the other members. The individual is motivated to seek

$$\max_t \; u(c-t, \frac{t}{T+t} \phi(T+t)).$$

This is achieved by

$$0 = - u_1 + u_2 \cdot [\frac{t\phi' - \phi}{T+t} - \frac{t\phi}{(T+t)^2}].$$

We must show that the bracket equals ϕ'.
Substitution for ϕ' from (27) yields

$$[\;] = \frac{1}{T+t} (\frac{t}{nt} \phi' + \phi - \frac{t}{T+t} \phi) = \frac{\phi}{T+t} \quad \text{since} \quad nt = T+t$$

$$= \frac{\phi}{T+t} = \frac{\phi}{nt} = \phi' \quad \text{by (27).} \qquad \text{Q.E.D.}$$

15.6 Consider now

<u>Case 3.</u> Social pressure: Each person must put in at least an average effort

$$t_i \geqq \frac{T}{n-1} = \tau$$

This is realized only when the = applies for every person i. Then

$$\ln (c-t_i) + d_i \ln (\frac{1}{n} \phi(T+t_i))$$

$$= \ln (c - \frac{T}{n-1}) + d_i \ln (\frac{1}{n} \phi(T \frac{n}{n-1}))$$

$$= \ln (c-\tau) + d_i \ln \frac{1}{n} \phi(n\tau)).$$

When preferences are equal,

$$d_i \equiv d,$$

social pressure is equivalent to proportional sharing. When not, then group pressure pushes τ up to the level preferred by the person with the least preference for leisure.

15.7 Finally compare the utility $\tilde{u}$ of partnerships with that of an organizer of a hierarchical organization. We assume that managing a hierarchical organization is always a full-time job, $t = 1$. Achieved utility is larger in managing a hierarchical organization when

$$\ln (c-1) + d \ln k$$

$$> (1+d) \ln \frac{c}{1+d} + d \ln gd, \quad \text{i.e., whenever}$$

$$\ln \frac{k}{g} > \frac{1+d}{d} \ln \frac{c}{1+d} + \ln d - \frac{1}{d} \ln (c-1).$$

Assuming $c = 2$ this simplifies

$$\ln \frac{k}{g} > \ln d + \frac{1+d}{d} \ln \frac{2}{1+d}.$$

The right-hand side is zero for $d = 1$. It increases to infinity as d goes to zero. Thus partnerships are attractive to persons whose preference for leisure is large relative to their preference for income.

16. The Economics of Hierarchy

We now turn to the question raised in the beginning of this chapter: if organization is advantageous at the simple level $R = 1$, what, if any, are the economic advantages of hierarchical organizations with more than one level of supervision, $R > 1$?

16.1 At the outset it is useful to distinguish between two styles of management: management by team work and management by delegation. The first combines the inputs of $x_1 = 1$ supervisor with $x_0 = x$ subordinates in a joint effort. It is described by the production function

$$F(x_0, x_1) = F(1, x)$$

and is assumed to be the applicable type in simple organizations.

In management by delegation each person acts on his/her own with a certain amount of supervision by his/her supervisor. In terms of the production function we have

$$x \quad \text{individuals producing} \quad F(1, \tfrac{1}{x})$$

for a total of

$$x\, F(1, \tfrac{1}{x}).$$

This will be written

$$f(x,1) \equiv xF(1, \frac{1}{x}) \qquad \text{or more generally}$$

$$f(x_0,x_1) \equiv x_0 F(1, \frac{x_1}{x_0})$$

By construction $f(x_0,x_1)$ is linear homogeneous.
The reader should verify that

$$f(\lambda x_0, \lambda x_1) = \lambda f(x_0,x_1)$$

Define now

$$\underset{x}{\text{Max}} \quad xF(1, \frac{1}{x}) - x = h \tag{1}$$

Then

$$m < h < k. \tag{2}$$

Proof: To show the right inequality observe that

$$x\, F(x, \frac{1}{x}) = x^{1-m}\, F(x,1) \qquad \text{by homogeneity of } F$$

$$< F(x,1) \quad \text{for} \quad x > 1 \quad m > 1$$

(Recall that the maximizing x is not less than $\frac{1}{v} > 1$.)
To prove the left-hand inequality write

$$x\, F(1, \frac{1}{x}) - x = x(1 + \frac{1}{x})^m\, F(\frac{1}{1+\frac{1}{x}}, \frac{1/x}{1+\frac{1}{x}}) - x \tag{3}$$

Now $\quad F(\frac{1}{1+\frac{1}{x}}, \frac{1/x}{1+\frac{1}{x}}) = 1$

for $\frac{1}{x+1} = v$ or $x = \frac{1}{v} - 1 = \hat{x}$ (say). Substituting in (3)

$$= \hat{x}\lfloor (1 + \frac{1}{\hat{x}})^m - 1 \rfloor$$

$$\geqq \hat{x}\lfloor 1 + \frac{m}{\hat{x}} - 1 \rfloor = m \qquad \text{using a Taylor expansion.}$$

Therefore

$$h = \underset{x}{\text{Max}} \; x\,F(1, \frac{1}{x}) - x > m. \qquad \text{Q.E.D.}$$

We have thus shown that m is a lower bound to k as well.

Example: Consider the Cobb-Douglas case

$$F(x_0, x_1) = (\alpha+\beta)^{\alpha+\beta} \alpha^{-\alpha} \beta^{-\beta} x_0^{\alpha} x_1^{\beta}$$

$$= m^m (m-\beta)^{\beta-m} \beta^{-\beta} x_0^{\alpha} x_1^{\beta}$$

$$f(x,1) = x\,F(1, \frac{1}{x})$$

$$= m^m (m-\beta)^{\beta-m} \beta^{-\beta} x^{1-\beta}$$

Then $\underset{x}{\text{Max}} \; f(x,1) - x$ yields

$$h = m^{\frac{m}{\beta}} (m-\beta)^{\frac{\beta-m}{m}} (1-\beta)^{\frac{1-\beta}{\beta}}$$

it may be shown (by L'Hopital's rule) that $h = m$ for $\beta = 0$ and h is an increasing function of β.

Our principal result is the following:

Theorem:

$$f(x_o, f(x_1, \dots f(x_{R-1} 1) \dots) \leqq x_o + hx_1 + \dots + h^{R-1} x_{R-1} + h^R \qquad (4)$$

and "=" if

and only if $x_r = \sigma(\sigma+h)^{R-r-1}$ $\quad r = 0, \dots, R-1$

Here σ denotes the maximizer of $f(x,1) - x$. This theorem will be called "The Theorem on Returns to Scale in Nested Organizational Production Function" for reasons explained below. The proof will be given for $R = 3$. Consider

$$f(x_o, f(x_1, f(x_2,1))) \leqq f(x_o, f(x_1, h + x_2))$$

using (1);

$$= (h+x_2)\ f(\frac{x_o}{h+x_2}, f(\frac{x_1}{h+x_2}), 1))$$

using linear homogeneity of f;

$$\leqq (h + x_2)\ f(\frac{x_o}{h+x_2}, h + \frac{x_1}{h+x_2})$$

using (1);

$$= (h+x_2)(h + \frac{x_1}{h+x_2}) \cdot f(\frac{\frac{x_o}{h+x_2}}{h + \frac{x_1}{h+x_2}}, 1)$$

using linear homogeneity;

$$\leqq (h+x_2)\ (h + \frac{x_1}{h+x_2})(h + \frac{\frac{x_o}{h+x_2}}{h + \frac{x_1}{h+x_2}})$$

using (1);

$$= h^3 + h^2 x_2 + hx_1 + x_o$$

The "=" is taken on when

$$x_2 = \sigma$$

$$\frac{x_1}{h+x_2} = \sigma \qquad (5)$$

$$\frac{\frac{x_o}{h+x_2}}{h + \frac{x_1}{h+x_2}} = \sigma$$

The first two equations imply

$$x_1 = (h+\sigma)\sigma \qquad (6)$$

Substituting in the last equation one has

$$x_o = (h+\sigma)^2 \sigma \qquad \text{Q.E.D.} \qquad (7)$$

This theorem can be interpreted as follows.

If a simple organization is advantageous and achieves the pay-off k with team management and h using management by delegation, then a multi-level organization is advantageous provided employees of rank r receive an opportunity wage h^r. The organization's profit or presidential payment is then h^R. Thus organizational advantage is escalated through the application of successive levels of management, but not at the rate k for team work, but at the lower rate h for management by delegation.

We conclude: whether a particular activity or industry is carried out in multi-level organizations rather than by simple organizations depends on the size of h relative to k. If h is sufficiently close to k then managers are willing to accept this as compensation rather than the full opportunity wage k that one might achieve on one's own. Of course, running a simple organization of one's own requires team work as well as various other things such as initiative, access to capital, etc., in other words, enterpreneurship rather than just managerial capability. If h is too small, then the activity is one for which only simple organizations will be found advantageous.

In addition, there will be activities where not even simple organizations are practical since $k \leqq 1$. At present there are only a few of these left, and these are found mainly in the fields of art and writing.

That managers receive a higher wage than operatives is consistent with long-run equilibrium and free entry provided we assume managerial talent to be limited and/or the effort level in management to exceed that required in operative work.

16.2 Consider now the possibility of control loss at a rate $\rho \leqq 1$. By definition the president's control is effective

$$y_R = x_R$$

but managerial control y_r at all levels $r < R$ is now reduced to

$$\rho y_r .$$

The managerial production function is then transformed to

$$\frac{y_r}{x_r} = \rho F(1, \frac{y_{r+1}}{x_r})$$

Substituting in the organizational production function

$$Q = f(x_o, \rho f(x_1, \ldots, \rho f(x_{R-1}, 1) \ldots))$$

The theorem on scale returns in nested production functions must now be modified as follows.

$$Q \leqq \rho^{R-1} h^R + \rho^{R-1} h^{R-1} x_{R-1} + \ldots + \rho h x_1 + x_o \qquad (8)$$

and

$$\text{"="} \quad \text{if} \quad x_r = \sigma[\rho^r (h+\sigma)]^{R-1-r}$$

The proof follows along the lines of the theorem and is left to the reader.

We interpret this result as follows. Loss of control can be corrected and compensated by making the following adjustments:

the rate of managerial pay is reduced from h^r to $(\rho h)^r$;

the president's compensation is reduced to

$$\rho^{R-1} h^R;$$

the span of control is reduced to

$$\rho(h+\sigma) \quad \text{for} \quad r = 1, \ldots, R-1$$

and is left unchanged for the president.

In fact, consider the total output under these conditions. Using (8)

$$Q = \rho^{R-1}h^{R} + \rho^{R-1}h^{R-1}\sigma + \rho^{R-2}h^{R-2}\ \rho\sigma(\sigma+h)$$

$$+ \ldots + \sigma\rho^{R-1}(\sigma+h)^{R-1}$$

$$= \rho^{R-1}h^{R} + \rho^{R-1}\sigma h^{R-1}[1 + \frac{\sigma+h}{h} + \ldots (\frac{\rho+h}{h})^{R-1}]$$

$$= \rho^{R-1}h^{R} + \sigma(\rho h)^{R-1}\ \frac{(\frac{\sigma+h}{h})^{R} - 1}{\frac{\sigma+h}{h} - 1}$$

$$= \rho^{R-1}h^{R} + \rho^{R-1}h^{R}[(\frac{\sigma+h}{h})^{R} - 1]$$

$$Q = \rho^{R-1}(\sigma+h)^{R} \qquad (9)$$

The total number of operatives is

$$n_o = \sigma\rho^{R-1}(\sigma+h)^{R-1} \qquad (10)$$

Per capita output of operatives is therefore

$$\frac{Q}{n_o} = \frac{\sigma+h}{\sigma} = \frac{1}{1 + \frac{h}{\sigma}} \qquad (11)$$

and this is independent of scale R and the control loss factor ρ. Therefore, with the appropriate adjustments of wages and spans of control, per capita output of workers is stabilized at a constant level. The organization operates with constant returns to scale.

Consider also total cost

$$W = \rho^{R-1}h^R + \sum_{r=0}^{R-1} h^r\rho^r \ \sigma[(\sigma+h)\rho]^{R-r-1}$$

$$= \rho^{R-1}h^R + \rho^{R-1} \ \sigma(\sigma+h)^{R-1} \sum_{r=0}^{R-1} \left(\frac{h}{\sigma+h}\right)^r$$

$$= \rho^{R-1}h^R + \rho^{R-1} \ \sigma(\sigma+h)^{R-1} \frac{1 - \left(\frac{h}{\sigma+h}\right)^R}{1 - \frac{h}{\sigma+h}}$$

$$= \rho^{R-1}h^R + \rho^{R-1}(\sigma+h)^R - \rho^{R-1}h^R$$

$$= \rho^{R-1}(\sigma+h)^R$$

and this is equal to total output (9).

Cost per unit output is, therefore, unity at all levels of organization and for individual effort. This is but another aspect of constant returns to scale. The surplus achieved through organization is paid out as wages to managers and presidential compensation.

16.3 In an alternative specification the heart of productive effort is assumed to reside in the president and his/her team. Their combined output is described by the original production function applicable to team work

$$F(x_{R-1}, 1)$$

The organizational production function is then modified to

$$Q = f(x_o, \rho f(x_1, \ldots, \rho f(x_{R-2}, F(x_{R-1}, 1) \ldots)$$

Applying the same argument as before one has

$$Q \leqq \rho f(x_{R-2}, k + x_{R-1})$$

$$= (k + x_{R-1})\ f(\frac{x_o}{k+x_{R-1}},\ \rho f(\frac{x_1}{k+x_{R-1}},\ \ldots)$$

from which eventually

$$Q \leqq k(\rho h)^{R-1} + (\rho h)^{R-1} x_{R-1} + \ldots + \rho h x_1 + x_o$$

$$= \quad \text{when} \quad x_r = s \cdot \rho(\sigma+h)^{R-r-1}$$

and s is the maximizer of

$$F(x,1) - x.$$

Here the president is rewarded by an exceptional salary increment

$$k > h,$$

which is due to top level team work. In fact, k is the return to a simple organization operating as a team.

16.4 In the long run the possibility of changes in output price p and wage w must be considered. We shall use a labor standard of value and hence keep the wage rate at unity, $w = 1$. Initially in equilibrium this was also the output price. The coefficients of the production function must then be such that

$$w = \underset{x_o+x_1=1}{\text{Max}}\ pF(x_o, x_1)$$

or in view of $w = 1$ $p = 1$

$$1 = \underset{x_o+x_1=1}{\text{Max}}\ F(x_o, x_1)$$

Thus in the case of Cobb-Douglas production functions it was shown in Section 15 that

$$1 = \max_{x_o + x_1 = 1} b x_o^{\alpha} x_1^{\beta}$$

implies

$$b = (\alpha+\beta)^{\alpha+\beta} \alpha^{-\alpha} \beta^{-\beta}$$

Assume that initially the management technology had constant returns to scale. There is then no incentive to organization since

$$k = \max_{x} F(x,1) - x = 1$$

so that an entrepreneur can earn no more than the opportunity wage of unity.

Suppose now that technical change generates a new production function with increasing returns to scale such that the production function is now homogeneous of degre $m > 1$. In the Cobb-Douglas case this means changes in α and/or β, but possibly also in the productivity coefficient a. Suppose now that

$$k = \max_{x} pF(x,1) - x > 1 \qquad p = 1$$

is sufficiently large to attract organizers $k \geqq k_o$. Then as a result of the increased supply, price p will fall. We show by an example that this price fall may drive out individual operators. This activity is then taken over by simple organizations.

Consider the Cobb-Douglas production function

$$y = b x_o^{\frac{3}{4}} x_1^{\frac{1}{2}}$$

An individual operator's earnings will be

$$w = \underset{x_o + x_1 = 1}{\text{Max}} \ pbx_o^{\frac{3}{4}} \ x_1^{\frac{1}{2}}$$

These earnings are maximized when

$$\frac{x_o}{x_1} = \frac{3}{2} \quad \frac{3/4}{x_1} = \frac{1/2}{x_1} \qquad \text{or} \quad x_o = \frac{3}{5} \quad x_1 = \frac{2}{5}$$

yielding

$$w = pb(\tfrac{3}{5})^{\frac{3}{4}} \cdot (\tfrac{2}{5})^{\frac{1}{2}}$$

$$= pb \ 0.5428$$

The boss of a simple organization earns

$$k = \underset{x}{\text{Max}} \ pb \ x^{\frac{3}{4}} - x$$

and this is achieved when

$$x = (\tfrac{3}{4} \ pb)^4$$

yielding

$$k = \tfrac{1}{4} \ (pb)^4 \ (\tfrac{3}{4})^3 - \tfrac{29}{256} \ (pb)^4$$

Now

$$w < 1 < k \qquad (12)$$

for all pb from

$$1.807 < pb < 1.844$$

In particular $pb = 1.84$ yields $k = 1.2089$.
The condition for (1) to hold with a Cobb-Douglas function is

$$pb \left(\frac{\alpha}{\alpha+\beta}\right)^{\alpha}\left(\frac{\beta}{\alpha+\beta}\right)^{\beta} < 1 < (pb)^{\frac{1}{1-\alpha}}\alpha^{\frac{\alpha}{1-\alpha}}(1-\alpha) \qquad (13)$$

This is satisfied whenever

$$\alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)} < pb < (\alpha+\beta)^{\alpha+\beta}\ \alpha^{-\alpha}\ \beta^{-\beta} \qquad (14)$$

Now it can be shown that the left-hand expression in (14) is strictly less than the right-hand expression for all

$$\alpha+\beta \geqq 1, \quad \alpha \geqq 0, \quad \beta \geqq 0$$

in fact

$$\ln\ [(\alpha+\beta)^{\alpha+\beta}] - \ln\ [\alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)}] = 0$$

for $\alpha+\beta = 1$

and this is an increasing function of α or β.

Therefore, the possibility of organizations driving out individuals always exists in a Cobb-Douglas technology with increasing returns.

On the other hand, since

$$k > h \qquad (2)$$

simple organizations will always yield a better return then can be earned by first line managers in a hierarchical organization.

Now first line managers must, in fact, accept lower rates of pay than individual entrepreneurs. In fact, they may be satisfied with a given $h > h_o$ when entrepreneurs are dissatisfied with a larger $k = k_o$.

Except for this possibility (which seems unlikely), as far as management technology is concerned, simple organizations can never be driven out of business by hierarchies. Even the automotive industry has many small firms specializing in various aspects of automobile production, modification and repair, while leaving mass production to hierarchical organizations.

When $m > 1$ but $\rho h < h_o$ is too small, then only simple organizations plus possibly individuals will survive in this activity.

To summarize we have the following possibilities:

$m < 1$		only part-time activity
$m = 1$		individuals only
$m > 1$	$1 < h < h_o$	rewards for organizations too small and rewards for managers too small: individuals only
$m > 1$ $\rho h > h_o$	$k < k_o$	rewards for simple organizations too small but adequate for managers: only hierarchies plus possibly individuals
$m > 1$	$k > k_o$ $\rho h < h_o$	rewards adequate for simple organizations, insufficient for managers: only simple organizations plus possibly individuals
$m > 1$	$k > k_o$ $\rho h > h_o$	both hierarchies and simple organizations, no individuals

We conclude with the following remark: The advantage of organization discussed here must be seen against a background of a competitive market economy, in which organizations appear as price

takers. The question of economic efficiency changes its nature entirely when organizations become large enough to act as monopolists or as agents of central planning. So does the role of individuals in organizations. By the nature of the control process, freedom of action is restricted in organizations. The only effective protection against enslavement by organizations is their competition for personnel in competitive labor markets.

Seen from this broader perspective, organizations are not only a source of inequality--of results, not necessarily of opportunities--they are also a potential threat to individual liberty. To enjoy a maximum of freedom, one must either be an independent individual operating on one's own (but possibly at an economic sacrifice) or must succeed in reaching the top position in an organization. However, "many are called, but few are chosen."

Supplement H: Qualification and Specialization

So far, all personnel was assumed to be equally qualified both in the operative and the managerial roles. Suppose now that there are two types of persons with qualification q and 1 respectively. These are the quantities of labor in efficiency units possessed by the two types of person. These quantities are assumed to apply both to managerial and operative work. We ask: should better qualified persons $(q > 1)$ be assigned to operative or to managerial work?

When used as operatives, the opportunity wage of qualified persons is q. What is their productivity in managerial jobs?

A qualified person used as manager contributes $x_1 = q$. In terms of the production function $F(x_0, x_1)$ with returns to scale of degree m, the surplus obtained by using a qualified person as manager equals

$$
\begin{aligned}
k_q &= \max_{\substack{x_0, x_1 \\ x_1 \leqq q}} F(x_0, x_1) - x_0 \\
&= \max_{\substack{x_0, x_1 \\ x_1 \leqq q}} q^m F\left(\frac{x_0}{q}, \frac{x_1}{q}\right) - x_0 \\
&= \max_{\substack{z_0, z_1 \\ z_1 < 1}} q^m F(z_0, z_1) - q \cdot z_0
\end{aligned}
$$

$$> q \max_{\substack{z_0, z_1 \\ z_0 < 1}} [F(z_0, z_1) - z_0]$$

$$= q \cdot k$$

by definition (14.9) of k.

This proves that a person of qualification q (in efficiency units) makes his/her largest contribution $k_q > q \cdot k > q$ as manager. The ablest should rise to the top.

References

Alchian, A. and H. Demsetz, 1972, "Production, Information Costs, and Economic Organizations," The American Economic Review, 62, pp. 777-795.

Arrow, Kenneth J., 1974, The Limits of Organization, New York: W.W. Norton.

Beckmann, M.J., 1960, "Returns to Scale in Business Administration," Quarterly Journal of Economics, 74, 464-471.

Beckmann, Martin 1977, "Management Production Functions and the Theory of the Firm," Journal of Economic Theory, 14,1, 1-18.

Beckmann, Martin J., 1978, Rank in Organizations, Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, New York.

Beckmann, Martin J., 1982, "Production Function for Organizations Doing Case Work, Management Science, Vol. 28, No. 10, October, 1159-1165.

Bell, Gerald D., 1967, "Determinants of the Span of Control," American Journal of Sociology, 73, 100-109.

Berge, Claude, 1976, Graphs and Hypergraphs, North-Holland.

Blau, Peter M. and W.R. Scott, 1963, Formal Organization: A Comparative Approach, London: Routledge and Kegan Paul.

Blau, Peter M., 1973, Bureaucracy, London: Institute for Economic Affairs.

The Budget of the United States, 1982, Washington, D.C., U.S. Government Printing Office.

Calvo, C.A., 1979, "Hierarchy, Ability, and Income Distribution," Journal of Political Economy, Vol. 87, No. 5.

Cox, D.R., 1967, Renewal Theory, London: Methuen.

Downs, A., 1967, Inside Bureaucracy, Boston: Little, Brown.

Dressler, Gary, 1980, Organization Theory, Englewood Cliffs: Prentice-Hall.

Drucker, Peter, 1954, The Practice of Management, New York: Harper.

Harary, Frank, 1971, Graph Theory, Addison-Wesley Publishing Company, Reading, Massachusetts.

Henn, R. and H.P. Kunzi, 1968, Einfuhrung in die Unternehmensforschung, I, II, Springer-Verlag, Berlin-Heidelberg-New York.

Kochen, M. and K.W. Deutsch, 1974, "A Note on Hierarchy and Coordination," Management Science, September, 106-114.

Koontz, Harold and Cyril P. O'Donnell, 1959, Principles of Management, New York: McGraw-Hill.

Lydall, H., 1968, The Structure of Earnings, Oxford: Clarendon Press.

Mackenzie, R.D., 1974, "Measuring a Person's Capacity for Interaction," Organizational Behavior, October, 149-169.

March, James G. and Herbert A. Simon, 1958, Organization, New York: Rand-McNally.

March, James G., 1965, Handbook of Organizations, Chicago: Rand-McNally.

Marschak, Jacob, 1952, "Teams and Organizations under Uncertainty," Cowles Commission Discussion Paper 2034.

Marschak, J., 1956, "Remarks on the Rationale of Leadership," Cowles Foundation Discussion Paper #6.

Marschak, J. and R. Radner, 1972, Economic Theory of Teams, New Haven: Yale University Press.

Marschak, J., 1977, "Efficient Organizational Design," Western Management Science Institute, Working Paper #273.

Marshall, A., 1891, Principles of Economics, London: Macmillan.

Mirrlees, J.A., 1976, "The Optimal Structure of Incentives and Authority Within an Organization," Bell Journal of Economics, Vol. 7, No. 1, 105-131.

Putterman, Louis, 1981a, "Labor Interdependence and Incentives in Commercial and Collective Enterprises," Working Paper No. 81-19, Department of Economics, Brown University.

Putterman, Louis, 1981b, "On Some Recent Explanations of Why Capital Hires Labor," Department of Economics, Brown University, Working Paper No. 80-25, Revised, December.

Radner, Roy and C.B. McGuire, 1972, Decision and Organization, Amsterdam: North-Holland.

Rosen, Sherwin, 1981, "Output, Income and Rank in Hierarchical Firms," Chicago: Economics Research Center NORC Discussion Paper #81-10.

Sato, R.S., 1981, Theory of Technical Change and Economic Invariance: Application of Lie Groups, Academic Press.

Sen, Amartya K., 1966, "Labor Allocation in a Cooperative Enterprise," Review of Economics Studies, 33, 361-71.

Simon, Herbert, 1965, Administrative Behavior; A Study of Decision-Making Processes in Administrative Organization, 2nd edition, New York: Free Press, 26-28.

Starbuck, William H., 1964, Mathematics and Organization Theory, Lafayette: Purdue University.

Starbuck, W.H., 1964, "Organizational Growth and Development," in J.G. March (ed.), Handbook of Organizations, Chicago: Rand-McNally & Co.

Tinbergen, Jan, 1964, Central Planning, New Haven: Yale University Press.

Wagner, Harvey M., 1969, Principles of Operations Research, Englewood Cliffs: Prentice-Hall.

Ward, Benjamin, 1958, "The Firm in Illyria: Market Syndicalism," American Economic Review, 48, 566-89.

Weber, Max, 1925, Wirtschaft und Gesellschaft, Grundriss der Sozialoekonomik, III, Abteilung, 2. Auflage, 1. Halbband, Tubingen: J.C.B. Mohr.

Williamson, Oliver E., 1975, Markets and Hierarchies, New York: The Free Press.

Author Index

ALCHIAN, A. 168
ARROW, Kenneth,J. 168
BECKMANN, Martin,J. 77,94,113, 123,135,168
BERGE,Claude 168
BELL,Gerald, D. 168
BLAU, Peter 168
CALVO, C.A. 94,135,168
COX,D.R. 103,168
DEMSETZ,H. 168
DEUTSCH,K.W. 94,169
DOWNS, A. 168
DRESSLER, Gary 168
DRUCKER, Peter 168
HARARY,Frank 21,22,28,29,54,168
HENN, Rudolf 169
KAUTILIYA 77
KOCHEN, M. 94,169
KOONTZ, Harold 35,169
KÜNZI, Hans 169
LYDALL,H. 121, 169
MACKENZIE,R.D. 169
MARCH, James 169
MARSCHAK, Jacob 169
MARSCHALL, Alfred 169
MC GUIRE, C. 169
MIRRLEES, J. 94,169
O'DONNELL, Cyril 35,169
PUTTERMAN, Louis 138,169
RADNER, Roy 169
ROSEN, Sherwin 135,170
SATO, Ryuzo 94,170
SCOTT, W.R. 168
SEN, Amartya 144, 170
SIMON, Herbert 169,170
STARBUCK, William 170
TINBERGEN, Jan 170
WAGNER, Harvey 96,103,170
WARD, Benjamin 139,170
WEBER, Max 170
WILLIAMSON, Oliver 91,93f,135,170

Subject Index

Antinode 23
Assignment 46

Branch 24f

Center 33f, 55f
Chain of command 6,14ff, 22, 33f, 56
critical(longest) 16f
Commune 144
Communication, loss of 89
Completion rate 95-104
time 95-98
Control 8ff
generation of preordering by 8ff
loss factor 158
loss of 89, 157
managerial 110
set 8,13
span: see span of control
Cost functions 115, 117f, 122f, 125
long run 123
medium run 118ff
short run 116ff

Degree 35f, 46, 53ff
Distance 14, 33f, 55,57
average 56f, 68f
total 56f
Divisibility requirement 40,49
Division of labor 135f

Eccentricity 33, 35
Elasticity of variable cost 123
Error rate 101f

Graph
  converse 19
  directed 5,19-22,25,27

Indegree 1,22,28
Information 89f

Labor cost (Wage bill) 70,73,77f, 83,159
  minimization of 70,73
  and scale 77ff, 88

Management
  by delegation 151,155
  by team work 151,155

Node 6,23,36

One-base 24
Organization
  acyclic 19,27
  acyclicity 7,19f,28
  advantage of 127, 132f, 151, 155, 164
  balanced 17f, 33, 41,44,49,56f, 68
  case working 94-108
  chart 6f,19,22,25f, 35,41
  connected
    unilaterally 20f
    weakly 20,25
  connectivity of. 7
  cyclic 7ff
  economics of 1,94
  height of 33, 86,115
  hierarchical 4,14,127, 135-140,146, 150f, 156,164

internal structure of 1
as ladder 9f,15,21,61,83f,86
nonregular 83
output of 128
regular 2,14,64f,67ff,73f,77,85,88,91
extended 3,79,91
quasi 3,81,83,88
simple 3,35,95,128f,131,151,155f,160,162,164
size of 37-40,46,49,62f,77f

Organizational
design 35,39,46ff
multiplier 38f
production function : see production function
slack 46, 49-52, 64, 115
amount of 50f
Outdegree 22
Outtree 21,25

Partnership 135-141, 144, 147f, 150
Pollacek-Kintchin Formula 103
Precedence 30f
ranking by 30f
President 4ff, 36
Pressure
group 150
social 149f
Production function 94, 97-100, 102f,105,109,116,119,120,131,137,139f 144,151, 159ff
Cobb-Douglas 92f, 100,113f, 116f, 119-123, 125, 132,142,153,161,163
management 109, 111, 113f, 129, 123,125,156
nested 154f, 157
organizational 94f, 112ff, 128, 130,157,159

Radius of a graph 55
Rank 1ff, 11-15, 60ff, 65, 80

assignment of 47f
average 60, 65ff
maximizing 60
minimizing 67ff, 115
communication aspects of 32ff
equivalence class of 12ff
ordering indiced by 12f

Returns
diminishing 99, 106,108,112,131
to scale 1,2, 129
constant 112, 129, 133, 158f, 161
decreasing 88, 129
increasing 129, 161, 163
and nested production function 154, 157

Semicycle 21, 25, 27f
Semipath 20,26
Span of Control 2 ff, 35f, 45, 53ff, 62,66,68,79,87,91,107f,120f, 136,146,157
average 41f, 44, 53f, 62, 121
constant 2, 36f, 57, 64, 92, 121
estimation of 41, 43f
maximal 36,64
optimal 35, 119, 121, 139, 141
Supervision 1ff, 4ff, 9f, 14f, 108f, 112, 130, 150f
amount of 37,39,46
graph theoretical model of 19
partnership vs. 135
structure of 2,32,36
Supervisory chain 16
relationship 5ff
as a directed graph 5ff
properties of 14f, 17,28,30,35
Support Structure 58f

Technical change 161

Unilateral component 24

Wage
  average 78,83f, 115
  bill: see labor cost
  scale 77, 82